

From: Bunner, Bridget
Sent: Monday, May 05, 2003 12:36 PM
To: STIC-Biotech/ChemLib
Subject: sequence search

Hi! I'd like to request a sequence search for case 09/894,912 (no pending search is required):

1. the amino acid sequence of SEQ ID NO: 13

Thanks!

Bridget Bunner

CRIG
Art Unit 1647
CM1-10D12
(703) 305-7148
mailbox 10B19

Point of Contact
P. Sheppard
Searcher: _____
Telephone number: (703) 308-4499
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: 5/7/03
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

AC	DT	DT	DT	FT	PROSITE; PS50032; TSPI; 1.
092132;	01-MAY-1999	(TREMBrel. 10, Created)	NON TER	224	
	01-JUN-2001	(TREMBrel. 17, Last annotation update)	SEQUENCE	224 AA;	224 MW;
		Thrombospondin type 1 domain.		97D26AD34CDBFB12	CRC64;
DE		R-SPONDIN.	SQ		
OS		Mus musculus (Mouse)			
OC		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
NCBI - TAXID=10090;			RN		
RR		SEQUENCE FROM N.A.			
RA		Kamatani T., Katsube K., Michikawa M., Yamada M., Mizusawa H.;			
RA		"R-spondin, a novel thrombospondin type 1 domain gene, expressed in			
RT		the dorsal neural tube.", to the EMBL/GenBank/DBJ databases.			
RI		Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.			
DR		EMBL: AB016768; BAA75640 1;			
DR		InterPro: IPR002174; Furin-like.			
DR		InterPro: IPR00884; TSPI.			
DR		PFam: PF00050; tsp_1; 1.			
DR		SMART; SM00261; FU; 2.			
DR		PROSITE; PS00092; TSPI; 1.			
DR		SEQUENCE 265 AA; 29331 MW; FFEBB864743F5963 CRC64;			
Query Match	42.5%	Score 644; DB 11; Length 265;			
Best Local Similarity	46.5%	Pred. No. 1, 2e-50;			
Matches	119;	Conservative			
		36; Mismatches			
		77; Indels			
		24; Gaps			
QY	6	LISWLFILNFMFYIGSONASRGRQRQRHPIVNSQCGGGATCSDYNGLSSPRFLFA 65	ID	026489	PRELIMINARY;
Db	11	VLSWTHIA-----VGSRG1-KGGRQRRIASAGSQCAGKGECECBVNGLSSPRFL 63	ID	026489	PRT; 1299 AA.
QY	66	LERIGMQLQIGVLSSQPSGGYGRYDPIKCTKCAD-CDTCFKNKNFTCKKCGFYHLG 124	AC	Q26489;	
Db	64	LENDRIVQVGGVLPSCPPGYFDRANDPKCKKICLHECEFSINFCKQKQALVHKG 123	AC	Q26489;	
QY	125	KCLDNCBPGELGAANNHMECVSIHCEVSEBWNNSPCTKKGKTCGKFRGTRTREIIOHP 184	DT	01-NOV-1996 (TREMBrel. 01, Created)	
Db	124	RCYPACBEGSTGAANSTMECCSACQCEMSEWSPPGCPSKKRLCGFRKGSEERTRVHLAP 183	DT	01-JUN-2002 (TREMBrel. 21, Last annotation update)	
QY	185	SAKGNLCPTPTNETRKCTVORKCOKGERGKKG---RERKKEPKNGKEAPDSKLE 240	DE	Endopeptidase Furin.	
Db	184	GGDHTTCSDTIETRCKTVRTPCPEGQKRKGCGQGRRENANRHPARKNSK--PRNSN 239	DR		
QY	241	SSKEIPORENRQQK 256	DR		
Db	240	-----RRIKGGQO 247	DR		
RESULT 5					
090GB2		PRELIMINARY;			
ID	090GB2	PRT; 224 AA.			
AC	090GB2;				
DT	01-MAY-2000 (TREMBrel. 13, Created)				
DT	01-OCT-2001 (TREMBrel. 18, Last sequence update)				
DR	01-DEC-2001 (TREMBrel. 19, Last annotation update)				
DR	0824F16.3 (Novel protein similar to mouse thrombospondin type 1 domain protein R-spondin) (Fragment).				
DR	D0824F16.3.				
OS	Hom sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Cetartiodactyla; Homidae; Homo.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Blakely S.;				
RA	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.				
EMBL; AL050125; CAB65783; 3;					
DR	InterPro; IPR002174; Furin-like.				
DR	InterPro; IPR000884; TSPI.				
DR	SMART; SM00261; FU; 2.				
DR	SMART; SM00209; TSPI; 1.				
QY	37	NVSOCGCGCATCSD-YNGCLSKCRPFALERIGMKQIGVQVLLSSCPSPCGYGRYDPIK 95	DR		
Db	795	SVCPCPAAMCATSERADGTCCTCEHHL-----VHDGTGKASCPSPHYET--EDDM 843	DR		
QY	96	CTKKAKADCTC--NIKFCFKKSFSYHLGKLDNPEGLEANNHMECVSIHCEVSE 153	DR		
Db	844	CAKCHESCDTCGCGPGETCVCYHPSPTVYALGRCVTCSPAYADKCKRKEC--HRCPG- 899	DR		

Qy 154 WPPWSPOTKKGKTCGKFKGKTRETRWEILOHPSAKNLCP--TNETRKC-TVORKKQK 209
 Qy 900 --CSTCT-----SAFCUSCPEKWEILKKKCMVPGSKCSA 933
 Qy 210 GRRGKKGCRERKRKRKPNK--GESKE--AIPSKSLESSKEIPE 247
 Qy 934 GEFAV--DOKCRNPACDSGYGENBGHCLTCPNPNLQDYKCP 977
 RESULT 7
 Q964D2 PRELIMINARY; PRT; 1101 AA.
 AC Q964D2; 01-DEC-2-2001 (TREMBREL. 19, Created)
 DT 01-DEC-2001 (TREMBREL. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBREL. 21, Last annotation update)
 DE Gal/GalNAc lectin IgG1.
 GN IgG1.
 OS Entamoeba histolytica.
 OC Eukaryota; Entamoebidae; Entamoeba.
 OX NCBI_TAXID=5759;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HML-IMSS;
 RX MEDLINE=21391835;
 RA Cheng X.J., Hughes M.A., Huston C.D., Loftus B., Gilchrist C.A.,
 Lockhart L.A., Ghosh S., Miller-Sims V., Mann B.J., Petri W.A. Jr.,
 RA Tachibana H.;
 RT "Intermediate Subunit of the Gal/GalNAc Lectin of Entamoeba
 histolytica Is a Member of a Gene Family Containing Multiple CXXC
 RT Sequence Motifs";
 RL Infect Immun. 69:892-898 (2001).
 DR AEMB; AAF331950; AAH92361.1; -.
 DR InterPro; IPR000563; EGF-like;
 DR PROSITE; PS01186; BGF_2; UNKNOWN_1.
 KW Lectin.
 SQ SEQUENCE 1101 AA; 119512 MW; CBB6F5CBDE656ABC CRC64;
 Query Match 11.7%; Score 177; DB 5; Length 1101;
 Best Local Similarity 26.6%; PkEd. No. 1e-07;
 Matches 55; Conservative 22; Mismatches 80; Indels 50; Gaps 10;
 AC Qy 44 CGCACCSDF--YNGCL--SC---KPRFLALE---RIGMKQIGVCLSSCPASYGTR 89
 Db 773 GTCOSLSDSKYPSGCKKFTSDNSRPGFYATECSDFGSGRSPYNSCTCTKSNY-PK 831
 Qy 90 YPDINKCTKKAQDPDTCKNKNCTK-----CKSGYIHLGKCLDNCPFGLA 136
 Db 832 EGECRKGCAKDKKATCSDKDFCLTADPLKVGSKCDGCKTGYMSNGC-----KHC 884
 Qy 137 NNHHTACVSLHCVSEINPAPCTKG-KTC--GFKGKTRETRWEILOHPSAKNLCP 193
 Db 885 TNHCSBCSSAECTVCESDTYKVISINGCNCVDPF 932
 Qy 194 TNETRKCTVORKKOKRGKGKGERK 220
 Db 933 TSPCTKGVKGKQDDEOPTGENSEKCK 939
 RESULT 8
 ID OBSZS2 PRELIMINARY; PRT; 1376 AA.
 AC OBSZS2; 01-JUN-2002 (TREMBREL. 21, Created)
 DT 01-JUN-2002 (TREMBREL. 21, Last sequence update)
 DE LD3182p.
 FN FUR2.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Sphyradoidea; Drosophilidae; Drosophila.
 OX NCBI_TAXID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 Amanatides P.G., Scherer S.E., Li P.W., Roskies R.A., Galle R.F.,
 George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 Brandon R.C., Rogers B.H., Blazek R.G., Champe M., Pfleiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
 Abril J.F., Agbayani A., An H.-J., Andreou-Pfannkoch C., Baldwin D.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Boleshkov S.,
 RA Borková D., Botchan M.R., Bouck J., Broksenin P., Brottier P.,
 RA Burtis K., Busam D.A., Butler H., Cadine E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Diez S.M.,
 RA Dobson K., Doop L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Fosler C., Gabril A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Correll J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Howland T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wai M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattai B., McIntosh M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Murry D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusslein-Volhard D.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shieh B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stickleton M., Strong R., Sun E.,
 RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*";
 RL *Science* 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ISO-1;
 RX MEDLINE:92301036; PubMed:1512259;
 RA Roobroek A.J., Creemers J.W., Pauli I.G., Kurzik-Dumke U., Rentrop M.,
 RA Gateff E.A., Leunissen J.A., de Ven W.J.,
 RT Cloning and functional expression of Dfurin2, a subtilisin-like
 RT proprotein processing enzyme of *Drosophila melanogaster* with multiple
 RT repeats of a cysteine motif";
 RL *J. Biol. Chem.* 267:17208-17215(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ISO-1;
 RX MEDLINE:95186060; PubMed:7880443;
 RA Roobroek A.J., Ayoubi T.A., Creemers J.W., Pauli I.G., Ven W.J.;
 RT "The Dfur2 gene of *Drosophila melanogaster*: genetic organization,
 RT expression during embryogenesis, and pro-protein processing activity
 RT of its translational product Dfurin2";
 RL *DNA Cell Biol.* 14:223-224(1995).
 DR EMBL: L31831; AA:48598_1; -;
 DR HSSP: Q39405; INPT:
 DR HSSP; S08_049; -;
 DR FlyBase: FBgn0004508; Fur2;
 DR InterPro: IPR000561; EGF-like;
 DR InterPro: IPR002174; Furin-like;
 DR InterPro: IPR000209; Peptidase_S8.
 DR InterPro: IPR002884; P-domain;
 DR Pfam: PF01483; P; 1.
 DR Pfam: PF00082; Peptidase_S8; 1.
 DR PRINTS: PRO0072; SUBTILISIN.
 DR Prodom: PDO00717; P domain; 1.
 DR SMART: SM00181; EG2; 1.
 DR SMART; SM00261; FU; 10.
 DR PROSITE; PS00136; SUBTILASE ASP; 1.
 DR PROSITE; PS00137; SUBTILASE HIS; 1.
 DR PROSITE; PS00137; SUBTILASE SER; 1.
 SQ SEQUENCE 1679 AA; 183368 MW; JF9E749FB021CF6 CRC64;

Query Match 11.6%; Score 176.5; DB 5; Length 1679;
 Best Local Similarity 28.0%; Pred. No. 1.7e-07; Matches 60; Conservativeness 24; Mismatches 77; Indels 53; Gaps 12;

QY 3 HRLLSWFLILPMIPEYIGSQASRRQRMRHNVNSCGCGCATSDV-NGCLSCXPR 61
 DB 1037 HLLHVID-LAVCLOPCPDGYFENS---RNRCTCVP---CECNCSAQDPEYCTCSHH 1086
 QY 62 LFFALERIGKQICVCLCSPCGYIGTRIPDINKCTKKAKADDTCP---NKNFCTYKCKSGF 119
 DB 1087 LVMHEH-----CYSACRPLDVT--EDNKAFCFHSPCATCNGPPTQDCTCCTCRR 1135
 QY 120 YLHLGKCLDNCPGEGLEANNHTMCYISIVHCBVSEWNPSPCTKKGKTCGFKRGTTFRV-R 178
 DB 1136 YAWQNKLISCPDPFYADKKLECM-----PCQEGCTC---TSNGVCS 1176

QY 179 EIQCHPSAKGNLCPPTNETRKCTVQRKK-CORGE 211
 DB 1177 ECIQWNT-----LNKRDKCTVSGSEGESE 1202

RESULT 10

ID 0964D1 PRELIMINARY; PRT; 1074 AA.
 AC 0964D1;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Gal/GalNAc lectin Ig12 (Fragment).

GN IGL2.

OS Entamoeba histolytica.

OC Eukaryota; Entamoebidae; Entamoeba.

OX NCBI_TaxID=5759;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M1-IMSS;
 RX MEDLINE:21391855; PubMed:11500668;
 RA Cheng X.J., Hughes M.A., Huston C.D., Loftus B., Gilchrist C.A., Jr.,
 RA Lockhart L.A., Ghosh S., Miller-Sims V., Mann B.J., Petri W.A., Jr.,
 RA Tachibana H.;
 RT "Intermedite Subunit of the Gal/GalNAc Lectin of Entamoeba histolytica Is a Member of a Gene Family Containing Multiple CXXC Sequence Motifs";
 RL Infect. Immun. 69:5892-5898(2001).
 DR EMBL: AP317951; AA:92362_1; -;
 DR InterPro: IPR000561; EGF-like.
 DR PROSITE; PS00186; EGG_2; UNKNOWN_1.
 KW Lectin.
 FT NON-TER 1
 FT NON-TER 1
 FT 1074 1074 AA; 116782 MW; 1552E2D714EB450F CRC64;

Query Match 11.6%; Score 176; DB 5; Length 1074;
 Best Local Similarity 25.6%; Pred. No. 1.3e-07; Matches 53; Conservativeness 23; Mismatches 81; Indels 50; Gaps 9;

QY 44 GGCATCSD---YNGC---LSCKRPLFALLE---RIGMKQIGVCLSCPSCYGR 89
 DB 757 GTCGCSDSLSKYPGKTTDTCNVDSRTGYIAYTCGDFSGSRSPYNSCTCIESNYY-PK 815

QY 90 YPDINKTCKKAQDCTPENKNE-----CTICKSGSFYHLGKCLDNCPGEGLEA 136

DB 816 EGEKNGCAKCDDKCATSDKDTCLTCTDPLKIGSKCDECKTGYMSNGC-----KPC 868

QY 137 NNHTBCVSVIHRGEVSEBNPNSPCTKKG-KTC---GFKRGTTETRVRB1IOHPSAKGNLCP 193

DB 869 TNHCSBCSSAECTCVCSDTYKVIGNGENACVDF-----YFDEJKGTCPC 916

QY 194 TNETRKCTVORKKQKGBERGKGRER 220

DB 917 TSPCKCQVKGKDCQEETGENSEK 943

RESULT 11

ID 091VK0 PRELIMINARY; PRT; 915 AA.

AC 091VK0;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)

DE Unknown (protein for MGC:18501).

GN GN

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10909;

RN [1]
 RP SEQUENCE FROM N.A.
 TISSUE=BREAST TUMOR;

RC

RA Strausberg R.
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC013068; AAH13088.1; -.
 MGD; MGI:97515; Pcsk5.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR000209; Peptidase_S8.
 DR InterPro; IPR002884; P_domain.
 DR Pfam; PF01483; P; 1.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR ProDom; PD000717; P_domain; 1.
 DR SMART; SM00181; EGF; 4.
 DR PROSITE; PS00136; SUBTILASE ASP; UNKNOWN_1.
 DR PROSITE; PS00137; SUBTILASE HIS; UNKNOWN_1.
 DR PROSITE; PS00138; SUBTILASE SER; UNKNOWN_1.
 SQ SEQUENCE 915 AA; 101486 MW; EE13FF8F7AFC7BEE CRC64;

Query Match 11.3%; Score 172; DB 11; Length 915;
 Best Local Similarity 28.4%; Pred. No. 2.5e-07;
 Matches 52; Conservative 26; Mismatches 53; Indels 52; Gaps 13;

Qy 36 PNVSO-GCQG-GATCSYNGCISCRPLFALERIGMKGQIGVCLLSSCPGYYGTRYPI 93
 Db 640 PBCSEVGCDGPGPBHCSD--CL---HYVYKL---NNTRICVASSCPGHY--HADK 685

Qy 94 NKCTKCKADCPCTCNK--NFCRKCKSGFYLUH--LGKCLDNCPGLEANNHMECVSIVHC 149
 Db 686 KCRKCKAPNCESCHGSHNQCLSKYGYFLNEBSCSVNQCPDG-----SY 731

Qy 150 ESEVNPNWPSCTKKGKTC-GFGRGTTTTRVRLIQHPSAKGNLCPPTNETRKTIVORKKQ 208
 Db 732 EDIKKVNGKCSENCKACIGHNCTE-----CKGGL---SLQGSRCSV--TCE 774

Qy 209 KGE 211
 Db 775 DQG 777

RESULT 12

035171 PRELIMINARY; PRT; 296 AA.
 AC 035171; 01-JAN-1998 (TREMBREL. 05, Created)
 DT 01-MAR-2002 (TREMBREL. 20, Last annotation update)
 DE PACE4 (Fragment).
 PSK6 OR PACE4.
 OS Mus musculus (Mouse)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Buterria; Rodentia; Sciurognathi; Muridae; Murine; Mus.
 OC NCB1_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ICR; TISSUE=LIVER;
 RA Nakayama K.;
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ICR; TISSUE=LIVER;
 RA Nakayama K.;
 RT 'PACE4 is ubiquitous endopeptidase that has similar but not identical substrate specificity to other Kex2-like processing endoproteases.';
 RT Blomed. Res. 15:363-370(1994).
 DR D50060; BAA0877.1; -.
 DR HSSP; Q99405; IMPT.
 DR MEROPS; S08.075; -.
 DR MGD; MGI:1102897; Pcsk6.
 DR InterPro; IPR00561; EGF-like.
 DR InterPro; IPR00214; Furin-like.
 DR InterPro; IPR000209; Peptidase_S8.
 DR InterPro; IPR002884; P_domain.
 DR InterPro; IPR00822; 2nF_C2H2.
 DR Pfam; PF01483; P; 1.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR Pfam; PF00723; SUBTILISIN.
 DR ProDom; PD000717; P_domain; 1.
 DR SMART; SM00181; EGF; 2.
 DR SMART; SM00001; EGF-like; 1.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR00214; Furin-like.
 DR InterPro; IPR00822; 2nF_C2H2.
 DR SMART; SM00181; EGF; 2.
 DR SMART; SM00261; FU; 5.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
 SQ SEQUENCE 296 AA; 32735 MW; 7796F39215AE231 CRC54;

Query Match 11.0%; Score 167.5; DB 11; Length 932;
 Best Local Similarity 24.4%; Pred. No. 6.5e-07;
 Matches 52; Conservative 25; Mismatches 78; Indels 61; Gaps 11;

Db	700	YFGDAARRCRR-----CHKGCTCTGGRSPAOCLSR-RGFY-----HHOETNT
Qy	77	CLSSCPGYYKTRYDPDINKCTCKADCDTCENK-NFCTKKSIGYLHIGKCLDNPEGLE 135
Db	743	CYVTLCPAGLYADESQRL-CLRCPHSOKCYDEPEKTCYKEPSLARGSCPDPCEPGTY 800
Qy	136	ANHHTMEC-----VSIHCEVS---EWNWPSCTKKGKTCGKFKGTRTR 176
Db	801	FDOSELVKPGCBEHHTCRTCWGPPSRBECIHCAKSFHQDWKCPVACGE----GF----- 848
Qy	177	VREIIOHPSAKGNLCPPTNETRKCTVORKKQQKGGER 212
Db	849	-----VPEENPGL--PHKVRCEENCLSCBGS 876
RESULT	14	
096EP4	096EP4	PRELIMINARY; PRT; 913 AA.
AC	Q96EP4;	
DT	01-DEC-2001 (Tremblrel. 19, Created)	
DT	01-JUN-2002 (Tremblrel. 19, Last sequence update)	
DE	Proprotein convertase subtilisin/kevin type 5.	
OS	Homo sapiens (Human)	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OC	NCBI_TaxID:9606;	
RN	{1}	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=KIDNEY;	
RA	Strausberg R.;	
RA	Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; B0012064; AAH12064.1; ..	
DR	MEOS4; S08_076; ..	
DR	InterPro; IPR000561; EGF-like.	
DR	InterPro; IPR002009; Peptidase_S8.	
DR	InterPro; IPR002884; P_domain.	
DR	PtMan; P01013; P; 1.	
DR	Pfam; PF00082; Peptidase_S8; 1.	
DR	ProDom; PD000717; P_domain; 1.	
DR	SMART; SM00181; EGF; 4.	
DR	PROSITE; PS00116; SUBTILASE ASP; UNKNOWN 1.	
DR	PROSITE; PS00117; SUBTILASE HIS; UNKNOWN 1.	
DR	PROSITE; PS00118; SUBTILASE SER; UNKNOWN 1.	
DR	SEQUENCE; R913 HA; 101649 NW; T2A2E63EFC49104BC CRC64;	
Qy	21	GESONAAGRQRQRQRQRQRHPPNYSQCGGGCATCSYNG---CLSCPKPLFALBRIGMKGIVL 78
Db	677	GHYHADK-KRCKCAPN-----CEBFGSHGDOCKSKYQFL-----NBTNSCV 721
Qy	79	SSCPGYYCTTRYDPDINKCTCKADCDTCENKNTCKKGFYVHLGKCLDNPEGLE 137
Db	722	THCPDGSYQDKK--NLRKKSSENCKTCBPHNCTECRDGSLQSSRCVSCEDGRYF 778
Qy	138	NHTE-----CVS-----IWHCESEW----NPSPCTK 163
Db	779	GDCOPCPCHRCCATCAGAGADGCINCTEGYFMDRCVOSCSISYYFDHSSENNGYKSKRC 838
Qy	164	GRKTC-----GRKGTE-----TRREIIOHPSAKGNLCPPTNETRKCTV 202
Db	839	DISCUTONGPGFKNTCTSPGYYLIDGMCQMAICKDATEESWAEGFCMLVKUNLC-- 896
Qy	203	ORKKQK 209
Qy	897	QKVKQO 903

RESULT 15
 018003 ID 018003 PRELIMINARY; PRT; 440 AA.
 AC 018003; (TREMBREL. 05, Created)
 DT 01-JAN-1998 (TREMBREL. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBREL. 19, Last annotation update)
 DE R17.3 protein.
 GN R17.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Chromadorea; Rhabditida; Rhabditidae; Peledorinae; Caenorhabditis; NCBITaxonID=6239;
 RP [1]
 SEQUENCE FROM N.A.
 RA Barlow K.;
 RL Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-990095613; PubMed=9851916;
 RX none;
 RT "Genome sequence of the nematode *C.elegans*: A platform for investigating biology";
 RL Science 282:2012-2018 (1998).
 EMBL: 232809; CAB07269; 1; -;
 DR InterPro: IPR001212; Somatomedin_B.
 DR InterPro: IPR000884; TSP1.
 PTM: PT000020; TSP1; 1.
 SMART: SM00209; TSP1; 1.
 PROSITE: PS00524; SOMATOMEDIN_B; 1.
 PROSITE: PS00092; TSP1; 1.
 SEQUENCE: PSS00092; TSP1; 1.
 SQ 44 AA. Mol. 5040 MW: 42820B834F263EE6 CRC64:

Db 242 KEIPPORENQQKKRKYQDOKS 265

RESULT 6

US-10-125-852-25

; Sequence 25, Application US/10125852

; Publication No. US20030120341

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL STEM CELL GROWTH FACTOR:

; FILE REFERENCE: HIS-43A

; CURRENT APPLICATION NUMBER: US/10/125, 852

; CURRENT FILING DATE: 2002-08-20

; PRIOR APPLICATION NUMBER: US 6/0/316, 368

; PRIOR FILING DATE: 2001-08-30

; PRIOR APPLICATION NUMBER: US 09/799, 451

; PRIOR FILING DATE: 2001-03-05

; NUMBER OF SEQ ID NOS: 25

; SOFTWARE: Patentin version 3.1

; SEQ ID NO: 25

; LENGTH: 292

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-10-125-852-25

Query Match 97.1%; Score 1472; DB 9; Length 292;

Best Local Similarity 100.0%; Pred. No. 1 lie-97;

Matches 264; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query 3 HRLISWLFILNFMVEIGSONASRGRORMPNSCGCGCATCSDYNGCLSKPRL 62

Db 2 HRLISWLFILNFMVEIGSONASRGRORMPNSCGCGCATCSDYNGCLSKPRL 61

Query 63 FFALEIRGMQKIGVCLSCPSGYGTRYPDINKCTCKKACDCTCFKNFCTKSGFYLH 122

Db 62 FFALEIRGMQKIGVCLSCPSGYGTRYPDINKCTCKKACDCTCFKNFCTKSGFYLH 121

Query 123 LGKCDNCPGLEANNHMECVSIVHCESEWNPWSCTKGKTCGKTRGTRVREI1Q 182

Db 122 LGKCDNCPGLEANNHMECVSIVHCESEWNPWSCTKGKTCGKTRGTRVREI1Q 181

Query 183 HPSAKGNLCPPTNETRCKTVORKKCGKGKGRERKKGKPKNGKESKEAIPDSKLESS 242

Db 182 HPSAKGNLCPPTNETRCKTVORKKCGKGKGRERKKGKPKNGKESKEAIPDSKLESS 241

Query 243 KEIPPORENQQKKRKYQDOKS 266

Db 242 KEIPPORENQQKKRKYQDOKS 265

RESULT 5

US-09-894-912A-26

; Sequence 26, Application US/09894912A

; Publication No. US20030044792A1

; GENERAL INFORMATION:

; APPLICANT: Tang et al.

; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL STEM CELL GROWTH

; FILE REFERENCE: 2810/3726A

; CURRENT APPLICATION NUMBER: US/09/894, 912A

; CURRENT FILING DATE: 2002-05-10

; PRIOR APPLICATION NUMBER: To be assigned

; PRIOR FILING DATE: 2001-04-05

; PRIOR APPLICATION NUMBER: 60/266, 614

; PRIOR FILING DATE: 2001-02-05

; PRIOR APPLICATION NUMBER: 60/215, 733

; PRIOR FILING DATE: 2000-06-28

; PRIOR APPLICATION NUMBER: 09/757, 562

; PRIOR FILING DATE: 2001-01-09

; PRIOR APPLICATION NUMBER: 09/543, 774

; PRIOR FILING DATE: 2000-04-05

; NUMBER OF SEQ ID NOS: 48

; SOFTWARE: Patentin version 3.0

; SEQ ID NO: 25

; LENGTH: 265

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-09-894-912A-26

Query Match 97.1%; Score 1472; DB 9; Length 265;

Best Local Similarity 100.0%; Pred. No. 9.9e-98;

Matches 264; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query 3 HRLISWLFILNFMVEIGSONASRGRORMPNSCGCGCATCSDYNGCLSKPRL 62

Db 2 HRLISWLFILNFMVEIGSONASRGRORMPNSCGCGCATCSDYNGCLSKPRL 61

Query 63 FFALEIRGMQKIGVCLSCPSGYGTRYPDINKCTCKKACDCTCFKNFCTKSGFYLH 122

Db 62 FFALEIRGMQKIGVCLSCPSGYGTRYPDINKCTCKKACDCTCFKNFCTKSGFYLH 121

Query 123 LGKCDNCPGLEANNHMECVSIVHCESEWNPWSCTKGKTCGKTRGTRVREI1Q 182

Db 122 LGKCDNCPGLEANNHMECVSIVHCESEWNPWSCTKGKTCGKTRGTRVREI1Q 181

Query 183 HPSAKGNLCPPTNETRCKTVORKKCGKGKGRERKKGKPKNGKESKEAIPDSKLESS 242

Db 182 HPSAKGNLCPPTNETRCKTVORKKCGKGKGRERKKGKPKNGKESKEAIPDSKLESS 241

Query 243 KEIPPORENQQKKRKYQDOKS 266

Db 242 KEIPPORENQQKKRKYQDOKS 265

RESULT 7

US-09-894-912A-48

; Sequence 48, Application US/09894912A

; Publication No. US20030044792A1

; GENERAL INFORMATION:

; APPLICANT: Tang et al.

; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL STEM CELL GROWTH

; FILE REFERENCE: 2810/3726A

; CURRENT APPLICATION NUMBER: US/09/894, 912A

; CURRENT FILING DATE: 2002-05-10

; PRIOR APPLICATION NUMBER: To be assigned

; PRIOR FILING DATE: 2001-04-05

; PRIOR APPLICATION NUMBER: 60/266, 614

; PRIOR FILING DATE: 2001-02-05

; PRIOR APPLICATION NUMBER: 60/215, 733

; PRIOR FILING DATE: 2000-06-28

; PRIOR APPLICATION NUMBER: 09/757, 562

; PRIOR FILING DATE: 2001-01-09

; PRIOR APPLICATION NUMBER: 09/543, 774

; PRIOR FILING DATE: 2000-04-05

; NUMBER OF SEQ ID NOS: 48

SOFTWARE: PatentIn version 3.0
 SEQ ID NO: 48
 LENGTH: 292
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-894-912a-48

Query Match 97.1%: Score 1472; DB 9; Length 292;
 Best Local Similarity 100.0%; Pred. No. 1.1e-97; Matches 264; Conservat. 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 3 HRLISWLFILNMEYIGSONASRRQRMRHPPVSCGCGCATSDYNGCLSKPRL 62
 Db 2 HRLISWLFILNMEYIGSONASRRQRMRHPPVSCGCGCATSDYNGCLSKPRL 61
 Qy 63 FPALARIGMKQIGVCLSSCPGYYGTRYPDINKCKKCADCDTCFNFCKCKSGFYIHL 122
 Db 62 FPALARIGMKQIGVCLSSCPGYYGTRYPDINKCKKCADCDTCFNFCKCKSGFYIHL 121
 Qy 123 LGKCLDNCPGLEANNHMECVSIVHCESEBNPWSPECTKKGKTFKGKRTETRREIQL 182
 Db 122 LGKCLDNCPGLEANNHMECVSIVHCESEBNPWSPECTKKGKTFKGKRTETRREIQL 181
 Qy 183 HPSAKNLCPPTNTRKCTVORKKCKGKGGERKKPKMGSEKEAPDSKLESS 242
 Db 182 HPSAKNLCPPTNTRKCTVORKKCKGKGGERKKPKMGSEKEAPDSKLESS 241
 Qy 243 KEIPEORENQOQQKRVQDKKS 266
 Db 242 KEIPEORENQOQQKRVQDKKS 265

RESULT 8

US-09-845-763-166
 Sequence 166: Application US/09745763
 Patent No. US200635941

GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth

McCoy, John M.
 LaVallie, Edward R.
 Collins-Racine, Lisa A.

Evans, Cheryl
 Mergerg, David
 Treacy, Maurice
 Spaulding, Vicki

TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES

NUMBER OF SEQUENCES: 219

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genetics Institute, Inc.

STREET: 87 CambridgePark Drive

CITY: Cambridge

STATE: MA

CITY: Cambridge

STATE: MA

COUNTRY: U.S.A.

ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US09/745,763

FILING DATE: 18-Jun-2000

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Springer, Suzanne A.

REGISTRATION NUMBER: 41,323

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 498-8284

TELEFAX: (617) 876-5851

INFORMATION FOR SEQ ID NO: 166:

SEQUENCE CHARACTERISTICS:

LENGTH: 292 amino acids

RESULT 9

US-09-894-912a-16
 Sequence 16: Application US/09894912A
 Publication No. US2003044792A1

GENERAL INFORMATION:

APPLICANT: Tang et al.

TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL STEM CELL GROWTH

FILE REFERENCE: 28110/3/260A

CURRENT APPLICATION NUMBER: US/09/894, 912A

CURRENT FILING DATE: 2002-05-10

PRIOR APPLICATION NUMBER: To be assigned

PRIOR FILING DATE: 2001-04-05

PRIOR APPLICATION NUMBER: 60/266, 614

PRIOR FILING DATE: 2001-02-05

PRIOR APPLICATION NUMBER: 60/215, 733

PRIOR FILING DATE: 2000-06-28

PRIOR APPLICATION NUMBER: 09/757, 562

PRIOR FILING DATE: 2001-01-09

PRIOR APPLICATION NUMBER: 09/543, 774

PRIOR FILING DATE: 2000-04-05

NUMBER OF SEQ ID NOS: 48

SOFTWARE: PatentIn version 3.0

SEQ ID NO: 16

TYPE: PRT

LENGTH: 251

ORGANISM: Homo sapiens

US-09-894-912a-16

Query Match 92.3%: Score 1400; DB 9; Length 251;
 Best Local Similarity 100.0%; Pred. No. 1.3e-92; Matches 251; Conservat. 0; Mismatches 0; Indels 0; Gaps 0;

Qy 23 QNASKRQRMRHPPVSCGCGCATSDYNGCLSKPRLFALARIGMKQIGVCLSSCP 82
 Db 1 QNASKRQRMRHPPVSCGCGCATSDYNGCLSKPRLFALARIGMKQIGVCLSSCP 60
 Qy 83 SGYYTRYPDINKCKKCADCDTCFNFCKCKSGFYIHLGKCLDNCPGLEANNHME 142
 Db 61 SGYYTRYPDINKCKKCADCDTCFNFCKCKSGFYIHLGKCLDNCPGLEANNHME 120

RESULT 10
 US-09-894-912A-32
 ; Sequence 32, Application US/09894912A
 ; Publication No. US20030044792A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang et al.
 ; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL STEM CELL GROWTH
 ; TITLE OF INVENTION: FACTOR-LIKE POLYPEPTIDES AND POLYNUCLEOTIDES
 ; FILE REFERENCE: 28110/37260A
 ; CURRENT APPLICATION NUMBER: US/09/894 912A
 ; CURRENT FILING DATE: 2002-05-10
 ; PRIOR APPLICATION NUMBER: To be assigned
 ; PRIOR FILING DATE: 2001-04-05
 ; PRIOR APPLICATION NUMBER: 60/266,614
 ; PRIOR FILING DATE: 2001-02-05
 ; PRIOR APPLICATION NUMBER: 60/215,733
 ; PRIOR FILING DATE: 2000-06-28
 ; PRIOR APPLICATION NUMBER: 09/757,562
 ; PRIOR FILING DATE: 2001-01-09
 ; PRIOR APPLICATION NUMBER: 09/543,774
 ; PRIOR FILING DATE: 2000-04-05
 ; NUMBER OF SEQ ID NOS: 48
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO: 32
 ; LENGTH: 279
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 ;
 US-09-894-912A-32
 Query Match 86.4%; Score 1310.5; DB 9; Length 279;
 Best Local Similarity 87.1%; Pred. No. 3.3e-86; Mismatches 18; Indels 7; Gaps 4;
 Matches 242; Conservative 11; Mismatches 18; Indels 7; Gaps 4;
 ;
 Qy 3 HRLRISWFLILNFMETIGSONASRRQRMRHMHNPSOCQGGCATSDYNGCLSKPRL 62
 ;
 Db 2 HRLRISWFLILNFMETIGSONASRRQRMRHMHNPSOCQGGCATSDYNGCLSKPRL 61
 ;
 Qy 63 FFALERIGMKQIGVCLSCPGYGYTRPDINKTKCKADCDTCFNKNFCTKCKSGFYLH 122
 ;
 Db 62 FFALERIGMKQIGVCLSCPGYGYTRPDINKTKCKADCDTCFNKNFCTKCKSGFYLH 121
 ;
 Qy 123 LGCLDNPEGLEANNHTMEVSIHCESEWNNWSPCTKKGKTCGKRGTRVREIIQ 182
 ;
 Db 122 LGCLDNPEGLEANNHTMEVSIHCESEWNNWSPCTKKGKTCGKRGTRVREIIQ 181
 ;
 Qy 183 HPSAKGNLCPNINETRCKTVORKKOKGERGKKGERKPK 226
 ;
 Db 182 HPSAKGNLCPNINETRCKTVORKKOKGERGKKGERKPK 225
 ;
 ;
 RESULT 11
 US-10-185-770-4
 ; Sequence 32, Application US/09894912A
 ; Publication No. US20030044792A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang et al.
 ; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL STEM CELL GROWTH
 ; TITLE OF INVENTION: FACTOR-LIKE POLYPEPTIDES AND POLYNUCLEOTIDES
 ; FILE REFERENCE: 28110/37260A
 ; CURRENT APPLICATION NUMBER: US/09/894 912A
 ; CURRENT FILING DATE: 2002-05-10
 ; PRIOR APPLICATION NUMBER: To be assigned
 ; PRIOR FILING DATE: 2001-04-05
 ; PRIOR APPLICATION NUMBER: 60/266,614
 ; PRIOR FILING DATE: 2001-02-05
 ; PRIOR APPLICATION NUMBER: 60/215,733
 ; PRIOR FILING DATE: 2000-06-28
 ; PRIOR APPLICATION NUMBER: 09/757,562
 ; PRIOR FILING DATE: 2001-01-09
 ; PRIOR APPLICATION NUMBER: 09/543,774
 ; PRIOR FILING DATE: 2000-04-05
 ; NUMBER OF SEQ ID NOS: 48
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO: 14
 ; LENGTH: 160
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ;
 US-09-894-912A-14
 Query Match 59.6%; Score 903; DB 9; Length 160;
 Best Local Similarity 100.0%; Pred. No. 1.9e-57; Mismatches 0; Indels 0; Gaps 0;
 Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 ;
 Qy 96 CTCKAKADCDTCFNKNFCTKCKSGFYLHKGCLDNPEGLEANNHTMEVSIHCESEWNNWSPCTKKGKTCGKRGTRVREIIQ 155
 ;
 Db 1 CTCKAKADCDTCFNKNFCTKCKSGFYLHKGCLDNPEGLEANNHTMEVSIHCESEWNNWSPCTKKGKTCGKRGTRVREIIQ 150
 ;
 ;
 RESULT 12
 US-09-894-912A-14
 ; Sequence 14, Application US/09894912A
 ; Publication No. US20030044792A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang et al.
 ; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL STEM CELL GROWTH
 ; TITLE OF INVENTION: FACTOR-LIKE POLYPEPTIDES AND POLYNUCLEOTIDES
 ; FILE REFERENCE: 28110/37260A
 ; CURRENT APPLICATION NUMBER: US/09/894 912A
 ; CURRENT FILING DATE: 2002-05-10
 ; PRIOR APPLICATION NUMBER: To be assigned
 ; PRIOR FILING DATE: 2001-04-05
 ; PRIOR APPLICATION NUMBER: 60/266,614
 ; PRIOR FILING DATE: 2001-02-05
 ; PRIOR APPLICATION NUMBER: 60/215,733
 ; PRIOR FILING DATE: 2000-06-28
 ; PRIOR APPLICATION NUMBER: 09/757,562
 ; PRIOR FILING DATE: 2001-01-09
 ; PRIOR APPLICATION NUMBER: 09/543,774
 ; PRIOR FILING DATE: 2000-04-05
 ; NUMBER OF SEQ ID NOS: 48
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO: 14
 ; LENGTH: 160
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ;
 US-09-894-912A-14
 Query Match 59.6%; Score 903; DB 9; Length 160;
 Best Local Similarity 100.0%; Pred. No. 1.9e-57; Mismatches 0; Indels 0; Gaps 0;
 Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 ;
 Qy 96 CTCKAKADCDTCFNKNFCTKCKSGFYLHKGCLDNPEGLEANNHTMEVSIHCESEWNNWSPCTKKGKTCGKRGTRVREIIQ 155
 ;
 Db 1 CTCKAKADCDTCFNKNFCTKCKSGFYLHKGCLDNPEGLEANNHTMEVSIHCESEWNNWSPCTKKGKTCGKRGTRVREIIQ 150
 ;
 ;

Qy 156 PWSPTCKGKTCGKGKGTRTVERVRELIQHPSAKONLCPPTNETRKCTVORKCOKGERGKK 215
 ; Software: PatentIn version 3.1
 ; Seq ID No: 21
 ; Length: 243
 ; Type: PRT
 ; Organism: Homo sapiens
 Db 61 PWSPCTKKGKTCGKGTRTVERVRELIQHPSAKONLCPPTNETRKCTVORKCOKGERGKK 120
 Qy 216 GRRKRKPKNPKGESEKEAIPDSKSESSKEIPQRENKQQQ 255
 Db 121 GRRKRKPKNPKGESEKEAIPDSKSESSKEIPQRENKQQQ 160
 RESULT 13
 US-10-125-852-18
 ; Sequence 18, Application US/10125852
 ; Publication No. US/0030032034A1
 ; General Information:
 ; APPLICANT: Tang, Y. Tom
 ; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL STEM CELL GROWTH FACTOR-L
 ; FILE REFERENCE: HYS-43A
 ; CURRENT APPLICATION NUMBER: US/10/125,852
 ; CURRENT FILING DATE: 2002-08-20
 ; PRIOR APPLICATION NUMBER: US 60/316,368
 ; PRIOR FILING DATE: 2001-08-30
 ; PRIOR APPLICATION NUMBER: US 09/799,451
 ; PRIOR FILING DATE: 2001-03-05
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO: 18
 ; LENGTH: 263
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-125-852-18
 Query Match 43.3%; Score 656; DB 9; Length 263;
 Best Local Similarity 45.3%; Pred. No. 1 2e-19; Mismatches 76; Indels 22; Gaps 4;
 Matches 117; Conservative 43; Mismatches 76; Indels 22; Gaps 4;
 Qy 6 LISWLFILNFMEMYIGQNSAQRGRMRMHPVNSQCGCATSDYNGCISCKPRLFFA 65
 Db 11 VLSWTHITI----SSRQIKGRKRRISAEGSQACKGELSEINGCLKSPKLFFI 63
 Qy 66 LERIGMKQIGVCLSSCPSPGYYGRYRDPINKCTKCAD-COTCEFKNFKCTKCSGFYLHG 124
 Db 64 LERNDIQRQVGVLSPSCPPGIFDARNPDWNKCTKCKIECEACFSHNFCTKCKEGLYIHKG 123
 Qy 125 KCDNCPEGEGLANNTMCEVSIHCESEWNPSCPTKAKTCKFGKGTRTVERVRELIQH 184
 Db 124 RCPKAPCPGSSAANGTMECSSPAQCEMSEWSPWGPEPSKQOLCCFRSEERTRVYLHAP 183
 Qy 185 SAKGNICPPNETRKCTVORKCOKGERGKGRRERKKENKG---ESKEAIPDSKLE 240
 Db 184 VGDHAQASDTRCTKTRVPCPEGOKRRKGOGRENANRNLARKESKEAGAGSR-- 240
 Qy 241 SSKEIPEORENKOOKKR 258
 Db 241 -----RRKGQOOooooo 251
 RESULT 14
 US-10-125-852-21
 ; Sequence 21, Application US/10125852
 ; Publication No. US/0030032034A1
 ; General Information:
 ; APPLICANT: Tang, Y. Tom
 ; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL STEM CELL GROWTH FACTOR-L
 ; FILE REFERENCE: HYS-43A
 ; CURRENT APPLICATION NUMBER: US/10/125,852
 ; CURRENT FILING DATE: 2002-08-20
 ; PRIOR APPLICATION NUMBER: US 60/316,368
 ; PRIOR FILING DATE: 2001-08-30
 ; PRIOR APPLICATION NUMBER: US 09/799,451
 ; PRIOR FILING DATE: 2001-03-05
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO: 24
 ; LENGTH: 265
 ; TYPE: PRT
 ; ORGANISM: MUS musculus
 ; US-10-125-852-24
 Query Match 42.5%; Score 644; DB 9; Length 265;
 Best Local Similarity 46.5%; Pred. No. 8.4e-39; Mismatches 77; Indels 24; Gaps 6;
 Matches 119; Conservative 36; Mismatches 77; Indels 24; Gaps 6;
 Qy 6 LISWLFILNFMEMYIGQNSAQRGRMRMHPVNSQCGCATSDYNGCISCKPRLFFA 65
 Db 11 VLSWTHIA----VGSRGI-KGRKRRISAEGSQACKGELSEINGCLKSPKLFFI 63
 Qy 66 LERIGMKQIGVCLSSCPSPGYYGRYRDPINKCTKCAD-COTCEFKNFKCTKCSGFYLHG 124
 Db 64 LERNDIQRQVGVLSPSCPPGIFDARNPDWNKCTKCKIECEACFSHNFCTKCKEGLYIHKG 123
 Qy 125 KCDNCPEGEGLANNTMCEVSIHCESEWNPSCPTKAKTCKFGKGTRTVERVRELIQH 184
 Db 124 RCPKAPCPGSTAATMCESSPAQCEMSEWSPWGPEPSKQOLCCFRSEERTRVYLHAP 183
 Qy 185 SAKGNICPPNETRKCTVORKCOKGERGKGRRERKKENKG---RERKRNPKGESEKEAIPDSKLE 240
 Db 184 GGDHTICSDTRCTKTRVPCPEGOKRRKGOGRENANRNLARKESKEAGAGSR-- 239
 Qy 241 SSKEIPEORENKOOKKR 256

Db 240 -----|:||||: RRHKGQQ 247

Search completed: May 6, 2003, 14:53:51
Job time : 23 secs

GenCore version 5.1.4.p5_4778
Copyright (c) 1993 - 2003 Compugen Ltd.

Om protein - protein search, using sw model

Run on: May 6, 2003, 14:50:58 ; Search time 15 Seconds

Title: US-09-894-912A-13
Perfect score: 1516
Sequence: 1 MGHRLILISMLPITLINFNEYI..... QQKERRQVDPDKQKSVSUTH 273

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 242574 seqs, 2942222 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA,*
1: /cgn2_6/ptodata/1/iaa/5B_COMB.pep: *
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep: *
3: /cgn2_6/ptodata/1/iaa/5A_COMB.pep: *
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep: *
5: /cgn2_6/ptodata/1/iaa/PCUS_COMB.pep: *

Pred No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	173.5	11.4	969	2	US-08-284-941-2		Sequence 2, Appli
2	173.5	11.4	969	2	US-08-447-642-2		Sequence 2, Appli
3	11.4	969	4	US-08-236-503-2			Sequence 2, Appli
4	173.5	11.4	969	5	PCT-US93-02147A-2		Sequence 2, Appli
5	162	10.7	799	2	US-08-525-940-23		Sequence 2, Appli
6	162	10.7	799	2	US-08-976-838-23		Sequence 2, Appli
7	162	10.7	881	2	US-08-525-940-21		Sequence 2, Appli
8	162	10.7	881	2	US-08-976-838-21		Sequence 2, Appli
9	162	10.7	915	2	US-08-525-940-18		Sequence 2, Appli
10	162	10.7	915	2	US-08-975-838-18		Sequence 2, Appli
11	162	10.7	915	4	US-09-214-555B-7		Sequence 2, Appli
12	162	10.7	915	4	US-09-214-555B-7		Sequence 2, Appli
13	161	10.6	288	1	US-08-368-645-15		Sequence 2, Appli
14	157.5	10.4	288	2	US-08-525-940-15		Sequence 2, Appli
15	157.5	10.4	288	2	US-08-976-838-15		Sequence 2, Appli
16	149	9.8	568	1	US-07-862-021B-14		Sequence 2, Appli
17	149	9.8	568	1	PCT-US93-0316A-14		Sequence 2, Appli
18	148.5	9.8	523	1	US-08-185-431-18		Sequence 2, Appli
19	148.5	9.8	523	4	US-08-899-232-3		Sequence 2, Appli
20	147.5	9.7	802	1	US-07-862-021B-12		Sequence 2, Appli
21	147.5	9.7	802	1	US-08-311-2881-12		Sequence 2, Appli
22	147.5	9.7	802	5	PCT-US93-0316A-12		Sequence 2, Appli
23	146	9.6	1062	1	US-08-537-210A-2		Sequence 2, Appli
24	146	9.6	1068	1	US-08-111-821-2		Sequence 2, Appli
25	146	9.6	2556	1	US-08-182-431-17		Sequence 2, Appli
26	146	9.6	2556	4	US-08-899-231-2		Sequence 2, Appli
27	9.5	807	1	US-08-313-2888-10		Sequence 2, Appli	

Sequence 10, Appli
Sequence 10, Appli
Sequence 103, Appli
Sequence 5, Appli
Sequence 4, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 18, Appli
Sequence 20, Appli
Sequence 20, Appli
Sequence 104, Appli
Sequence 2, Appli
Sequence 7, Appli
Sequence 16, Appli
Sequence 19, Appli
Sequence 19, Appli
Sequence 1, Appli

RESULT 1
US-08-284-941-2
; Sequence 2, Application US/08284941
; Patent No. 5663756

GENERAL INFORMATION:
APPLICANT: BARR, PHILIP J
APPLICANT: KIEFER, MICHAEL C
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PACE 4 AND
TITLE OF INVENTION: PACE 4.1 GENE AND POLYPEPTIDES IN CELLS
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: COOLEY GODWARD CASTRO HUDDLESON & TATUM
STREET: FIVE PALO ALTO SQUARE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94306

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/284,941
FILING DATE: 2 August 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: NEELEY PH. D., RICHARD L.
REGISTRATION NUMBER: 30092
REFERENCE/DOCKET NUMBER: CHIR-009/01US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 843-5070
TELEFAX: (415) 857-0653
TELEX: 380816 COOLEY PA

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 969 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-284-941-2

Query Match 11.4%; Score 173.5; DB 2; Length 969;
Best Local Similarity 28.8%; Pred. No. 8e-07;
Matches 44; Conservative 16; Mismatches 54; Indels 39; Gaps 7;
Qy 19 YFQGDTAAARRCRR-----CHKGGETCSSLRAATQCLSCR-RGFY----HHQEMNT 779
Db 737 YFQGDTAAARRCRR-----CHKGGETCSSLRAATQCLSCR-RGFY----HHQEMNT 779

RESULT 7
 US-08-925-940-21
 Sequence 21, Application US/08525940
 Patent No. 586651
 GENERAL INFORMATION:
 APPLICANT: Franzusoff, Alex
 ADDRESS: Sheridan Ross & McIntosh,
 STREET: 1700 Lincoln Street, Suite 3500
 CITY: Denver
 STATE: Colorado
 COUNTRY: U.S.A.
 ZIP: 80203
 COMPUTER READABLE FORM:
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patient Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/525,940
 FILING DATE:
 CLASSIFICATION: 514
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 08/368,852
 FILING DATE: 01-JAN-1995
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 08/088,322
 FILING DATE: 07-JUL-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Connell, Gary J.
 REGISTRATION NUMBER: 32,020
 REFERENCE/DOCKET NUMBER: 2848-11-C2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (303) 863-9700
 TELEFAX: (303) 863-0223
 INFORMATION FOR SEQ ID NO: 21:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 881 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-925-940-21

Query Match 10.7%; Score 162; DB 2; Length 881;
 Best Local Similarity 22.3%; Pred. No. 6.9e-06; Gaps 12;
 Matches 55; Conservative 33; Mismatches 81; Indels 78; Gaps 12;

RESULT 7
 US-08-925-940-21
 Sequence 21, Application US/08525940
 Patent No. 586651
 GENERAL INFORMATION:
 APPLICANT: Franzusoff, Alex
 ADDRESS: Sheridan Ross & McIntosh,
 STREET: 1700 Lincoln Street, Suite 3500
 CITY: Denver
 STATE: Colorado
 COUNTRY: U.S.A.
 ZIP: 80203
 COMPUTER READABLE FORM:
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patient Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/976,838
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Connell, Gary J.
 REGISTRATION NUMBER: 32,020
 REFERENCE/DOCKET NUMBER: 2848-11-C2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (303) 863-9700
 TELEFAX: (303) 863-0223
 INFORMATION FOR SEQ ID NO: 21:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 881 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-925-938-21

Query Match 10.7%; Score 162; DB 2; Length 881;
 Best Local Similarity 22.3%; Pred. No. 6.9e-06; Gaps 12;
 Matches 55; Conservative 33; Mismatches 81; Indels 78; Gaps 12;

RESULT 8
 US-08-976-938-21
 Sequence 21, Application US/08976838
 Patent No. 5891259
 GENERAL INFORMATION:
 APPLICANT: Franzusoff, Alex
 ADDRESS: Sheridan Ross & McIntosh,
 STREET: 1700 Lincoln St., Suite 3500
 CITY: Denver
 STATE: Colorado
 COUNTRY: U.S.A.
 ZIP: 80203
 COMPUTER READABLE FORM:
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patient In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/976,838
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Connell, Gary J.
 REGISTRATION NUMBER: 32,020
 REFERENCE/DOCKET NUMBER: 2848-11-C2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (303) 863-9700
 TELEFAX: (303) 863-0223
 INFORMATION FOR SEQ ID NO: 21:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 881 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-976-838-21

Query Match 10.7%; Score 162; DB 2; Length 881;
 Best Local Similarity 22.3%; Pred. No. 6.9e-06; Gaps 12;
 Matches 55; Conservative 33; Mismatches 81; Indels 78; Gaps 12;

RESULT 8
 US-08-976-938-21
 Sequence 21, Application US/08976838
 Patent No. 5891259
 GENERAL INFORMATION:
 APPLICANT: Franzusoff, Alex
 ADDRESS: Sheridan Ross & McIntosh,
 STREET: 1700 Lincoln St., Suite 3500
 CITY: Denver
 STATE: Colorado
 COUNTRY: U.S.A.
 ZIP: 80203
 COMPUTER READABLE FORM:
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patient In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/976,838
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Connell, Gary J.
 REGISTRATION NUMBER: 32,020
 REFERENCE/DOCKET NUMBER: 2848-11-C2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (303) 863-9700
 TELEFAX: (303) 863-0223
 INFORMATION FOR SEQ ID NO: 21:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 881 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-976-838-21

Query Match 10.7%; Score 162; DB 2; Length 881;
 Best Local Similarity 22.3%; Pred. No. 6.9e-06; Gaps 12;
 Matches 55; Conservative 33; Mismatches 81; Indels 78; Gaps 12;

RESULT 11
 US-09-214-555B-2
 ; Sequence 2, Application US/09214555B
 ; Patent No. 6380171
 ; GENERAL INFORMATION:
 ; APPLICANT: INSTITUT DE RECHERCHE CLINIQUE DE MONTREAL
 ; TITLE OF INVENTION: PRO-PROTEIN CONVERTING ENZYME
 ; FILE REFERENCE: PRO-PROTEIN CONVERTER ENZ
 ; CURRENT APPLICATION NUMBER: US/09/214, 555B
 ; CURRENT FILING DATE: 1995-01-04
 ; PRIORITY APPLICATION NUMBER: 60/021, 008
 ; PRIORITY FILING DATE: 1996-07-26
 ; PRIORITY APPLICATION NUMBER: 2, 203, 745
 ; PRIORITY FILING DATE: 1997-04-25
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO: 2
 ; LENGTH: 915
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-214-555B-2

Query Match Similarity 10.7%; Score 162; DB 4; Length 915;
 Best Local Similarity 22.3%; Pred. No. 7, 1e-06; Mismatches 81; Indels 78; Gaps 12;
 Matches 55; Conservative 33; MisMatches 81; Indels 78; Gaps 12;

Qy 21 GSONASRGRQRQRMMRPPNSQCGGCGATSDYNG---CLUCKPRLFALERIGMKGQIGVCL 78
 Db 679 GHYHADK-KGRKRCAPN-----CESCGSHQDQMSCKGYFL-----NEETNSCV 723
 Qy 79 SSCPG-YGTYRPDINCKTKKAKADCTCENKINFCTKCKSGFLHKGKCLDNCPEGLEAN 137
 Db 724 THCPDPSQYDPTKK---NLCKRKSSENKTCTEFNCCTEGRDGSLQSRSVSCEDGRFN 780
 Qy 138 NHMB-----CVS-----IVPCESEW-----NPNSPCTKK 163
 Db 781 GODCOPCHARCATCAGAGADGCGINCTEGYPMEDORGCVQCSISYYFDHSSENGYKSKKC 840
 Qy 164 GKTC-----GFKRGTE-----TRVREI1QHPSAKGNLCPPTNTRKCTV 202
 Db 841 DISCLTCNGPGFKNTCPSGQYLLDLMGCOMGAIKDATEESWAEGFCMLVKNL--- 898
 Qy 203 ORKCKQK 209
 Db 899 QRKVLUQQ 905

RESULT 12
 US-09-214-555B-7
 ; Sequence 7, Application US/09214555B
 ; Patent No. 6380171
 ; GENERAL INFORMATION:
 ; APPLICANT: INSTITUT DE RECHERCHE CLINIQUE DE MONTREAL
 ; TITLE OF INVENTION: PRO-PROTEIN CONVERTING ENZYME
 ; FILE REFERENCE: PRO-PROTEIN CONVERTER ENZ
 ; CURRENT APPLICATION NUMBER: US/09/214, 555B
 ; CURRENT FILING DATE: 1995-01-04
 ; PRIORITY APPLICATION NUMBER: 60/021, 008
 ; PRIORITY FILING DATE: 1996-07-26
 ; PRIORITY APPLICATION NUMBER: 2, 203, 745
 ; PRIORITY FILING DATE: 1997-04-25
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO: 7
 ; LENGTH: 915
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-214-555B-7

Query Match Similarity 10.7%; Score 162; DB 4; Length 915;
 Best Local Similarity 22.3%; Pred. No. 7, 1e-06; Mismatches 81; Indels 78; Gaps 12;
 Matches 55; Conservative 33; MisMatches 81; Indels 78; Gaps 12;

Qy 21 GSONASRGRQRQRMMRPPNSQCGGCGATSDYNG---CLUCKPRLFALERIGMKGQIGVCL 78
 Db 679 GHYHADK-KGRKRCAPN-----CESCGSHQDQMSCKGYFL-----NEETNSCV 723
 Qy 79 SSCPG-YGTYRPDINCKTKKAKADCTCENKINFCTKCKSGFLHKGKCLDNCPEGLEAN 137
 Db 724 THCPDPSQYDPTKK---NLCKRKSSENKTCTEFNCCTEGRDGSLQSRSVSCEDGRFN 780
 Qy 138 NHMB-----CVS-----IVPCESEW-----NPNSPCTKK 163
 Db 781 GODCOPCHARCATCAGAGADGCGINCTEGYPMEDORGCVQCSISYYFDHSSENGYKSKKC 840
 Qy 164 GKTC-----GFKRGTE-----TRVREI1QHPSAKGNLCPPTNTRKCTV 202
 Db 841 DISCLTCNGPGFKNTCPSGQYLLDLMGCOMGAIKDATEESWAEGFCMLVKNL--- 898
 Qy 203 ORKCKQK 209
 Db 899 QRKVLUQQ 905

RESULT 13
 US-09-168-052-15
 ; Sequence 15, Application US/08368852
 ; Patent No. 591183
 ; GENERAL INFORMATION:
 ; APPLICANT: Frannzsoft, Alex
 ; TITLE OF INVENTION: CD4+ T-LYMPHOCYTE PROTEASES AND GENES
 ; NUMBER OF SEQUENCES: 15
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSE: Sheridan Ross & McIntosh
 ; STREET: 1700 Lincoln Street, Suite 3500
 ; CITY: Denver
 ; STATE: CO
 ; COUNTRY: U. S. A.
 ; ZIP: 80203
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/368, 852
 ; FILING DATE: 05-JAN-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Veriser, Carol Takington
 ; REGISTRATION NUMBER: 37,459
 ; REFERENCE/DOCKET NUMBER: 2848-11
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 303/863-9700
 ; TELEFAX: 303/863-0223
 ; INFORMATION FOR SEQ ID NO: 15:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 288 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-368-852-15

Query Match Similarity 10.6%; Score 161; DB 1; Length 288;
 Best Local Similarity 28.2%; Pred. No. 1e-06; Gaps 11;
 Matches 46; Conservative 25; MisMatches 52; Indels 40; Gaps 11;

Qy 36 PNVSO-GCQG-GCATCSYNGGGLSCKPRLFALERIGMKGQIGVCLSSCPSPSYGTYRPDI 93

Db 13 PECSEVCGCDPGPPHCDN---CL---HYVYKL---NNTRICVSSCPGHY---HADK 58
 Qy 94 NKCTKCKADCDTCFNK---NFCTKCKSGFYLH---LGKCLDNCPGEGLEANNHMECVSIHIC 149
 Db 59 KCRKCAPNCESCGSHGDOQMSCKYGYFLNEETNSCVTHCPDGSYQDTKKNLC---- 112
 Qy 150 EVSENNPMSPTCKGKIC-GFRGRTTRVREBIGHPSAKGNLC 191
 Db 113 -----RKSENKTC-TEFHXCTECR----DGLSLQGSRC 142

RESULT 14
 US-08-525-940-15
 Sequence 15, Application US/0852-940
 Patent No. 5666351
 GENERAL INFORMATION:
 APPLICANT: Franzusoff, Alex
 APPLICANT: Miranda, Luis R.
 APPLICANT: Wolf, Joseph R.
 TITLE OF INVENTION: CD4+ T-LYMPHOCYTE PROTEASES AND GENES
 TITLE OF INVENTION: ENCODING SAID PROTEASES
 NUMBER OF SEQUENCES: 25
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sheridan Ross & McIntosh
 STREET: 1700 Lincoln Street, Suite 3500
 CITY: Denver
 STATE: Colorado
 COUNTRY: U.S.A.
 ZIP: 80203

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/525,940
 FILING DATE:
 CLASSIFICATION: 514

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/168,852
 FILING DATE: 01-JAN-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/088,322
 FILING DATE: 07-JUL-1993

ATTORNEY/AGENT INFORMATION:
 NAME: Connal, Gary J.
 REFERENCE/DOCKET NUMBER: 2148-11-CL
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (303) 863-9700
 TELEFAX: (303) 863-0223

INFORMATION FOR SEQ ID NO: 15:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 288 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-525-940-15

Query Match Score 10.4%; Score 157.5; DB 2; Length 288;
 Best Local Similarity 26.6%; Pred. No. 4.7e-06; Pairs 49; Mismatches 49; Conservative 30; Indels 41; Gaps 12; Matches 49; Conservative 30; Mismatches 64; Indels 41; Gaps 12;

Db 13 PECSEVCGCDPGDPDHCDN---CL---HYVYKL---NNTRICVSSCPGHY---HADK 58
 Qy 36 PNVSQ-GCG-GATCSDYNGCLSKPRIFFALEBRIGKQIGVCLSCPGYGYTRYPDI 93
 Db 13 PECSEVCGCDPGDPDHCDN---CL---HYVYKL---NNTRICVSSCPGHY---HADK 58
 Qy 94 NKCTKCKADCDTCFNK---NFCTKCKSGFYLH---LGKCLDNCPGEGLEANNHMECVSIHIC 149
 Db 59 KCRKCAPNCESCGSHGDOQMSCKYGYFLNEETNSCVTHCPDGSYQDTKKNLC---- 112
 Qy 150 EVSENNPMSPTCKGKIC-GFRGRTTRVREBIGHPSAKGNLC 208
 Db 159 PCHR 162

Search completed: May 6, 2003, 14:53:24
 Job time : 17 secs

CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sb-sib.ch).
CC
CC TISSUE=Embryo; PubMed=8367492;
CC DR EMBL; L31894; AAM61987.1; -.
CC DR HSSP; Q99405; IMPR.
CC DR MEROPS; S08.075; -.
CC DR InterPro; IPR000561; EGF-like.
CC DR InterPro; IPR002174; Furin-like.
CC DR InterPro; IPR002884; P_domain.
CC DR InterPro; IPR00209; Peptidase_S8.
CC DR Pfam; PF00082; Peptidase_S8; 1.
CC DR Prints; PR00723; SUBTILISIN.
CC DR ProDom; PD00717; P_domain; 1.
CC DR SMART; SM00131; EGF; 1.
CC DR SMART; SM0261; FU; 5.
CC DR PROSITE; PS00136; SUBTILASE_ASPP; 1.
CC DR PROSITE; PS00137; SUBTILASE_HSS; 1.
CC DR PROSITE; PS00138; SUBTILASE_S8R; 1.
CC KW Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal; Calcium;
CC Cleavage on pair of basic residues; Repeat.
CC SIGNAL 1
CC PROPEP 46 132 POTENTIAL.
CC PT CHAIN 133 937 PAIRED BASIC AMINO ACID CLEAVING ENZYME
CC PT DOMAIN 133 454 CATALYTIC.
CC PT DOMAIN 477 615 HOMO_B.
CC PT DOMAIN 680 937 CYS-RICH MOTIF (CRM) REGION.
CC PT SITE 132 133 -CLEAVAGE (AUTO-).
CC PT SITE 534 536 CELL ATTACHMENT SITE (POTENTIAL).
CC PT ACT SITE 186 186 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC PT ACT SITE 227 227 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC PT ACT SITE 401 401 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC PT CARBOHYD 240 240 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC PT CARBOHYD 882 882 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC PT CARBOHYD 900 900 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 937 AA; 104053 MW; F3865557C3705C8 CRC64;

Query Match 10.3%; Score 156.5; DB 1; Length 937;
Best Local Similarity 23.6%; Pred. No. 0.0005; DB 1;
Matches 51; Conservative 24; Mismatches 80; Indels 61; Gaps 11;

Oy 19 VIGQNSAQRQRQRHPPNQGCCGATCSTYN--GCLKPRPFALERIOMKQIGV 76
Db 705 YFGDTAARRCCR-----CHKGCTCTGRSPQFCSCR-RGPFY----HHOEFTNT 747

Oy 77 CLSSCPGSGYGYTRPDINKCKCKADCTCFNK-NFCTKCKSGFVHLHGKCLDNCPPEGL 135
Db 748 CYTICPAGLYADESORL--CLRCHBSCOKVDEPEKSTVCKEGFSLIARSCIDPCBEGTY 805
Oy 136 ANNHTMEC-----VSIVRCEVs---EWNPNWSPCTKKGKTCFGKRGTR 176
Db 806 FDSELRIGECHTCRITCVGSPSREBCIHCAKSFPFQWVKCVPACGB-----GF----- 853

Oy 177 VREIIOHPSAKGNLCPPTNTRKTCVORKKQKGER 212
Db 854 -----YPEPEMGLL--PHKVCCRQDENCLSCGSSR 881

RESULT 9
FSPO_XENLA STANDARD; PRT; 803 AA.
ID 1
AC P3547;
DT 01-JUN-1994 (Rel. 2.9, Created)
DT 01-JUN-1994 (Rel. 2.9, Last sequence update)
DT 15-JUL-1999 (Rel. 3.8, Last annotation update)
DE F-spondin precursor.

OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC Xenopidae; Xenopus.
OX NCBI TaxID=8355;
RN 11
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX
RA Ruiz I Altaba A.; Cox C.; Jeessel T.M.; Klar A.;
RT "Ectopic neural expression of a floor plate marker in frog embryos
injected with the midline transcription factor Pintallavis.";
RT Proc. Natl. Acad. Sci. U.S.A. 90:8268-8272(1993).
RL
CC FUNCTION: PROMOTES THE ATTACHMENT OF SPINAL CORD AND SENSORY
CC NEURON CELLS AND THE OUTGROWTH OF NEURITES IN VITRO. MAY
CC CONTRIBUTE TO THE GROWTH AND GUIDANCE OF AXONS IN BOTH THE SPINAL
CC CORD AND THE PNS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN THE FLOOR PLATE.
CC -!- SIMILARITY: CONTAINS 6 TSP TYPE-1 DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sb-sib.ch).
CC
CC DR L09123; AAM19105.1; -.
CC DR PIR; A47723; A47723.
CC DR InterPro; IPR002861; Reeler.
CC DR SMART; SM00269; TSP1; 6.
CC DR PROSITE; PS50092; TSP1; 6.
CC KW Glycoprotein; Signal; Repeat; Cell adhesion.
CC FT SIGNAL 1
CC FT DOMAIN 665 716 TSP TYPE-1 5.
CC FT DOMAIN 751 803 TSP TYPE-1 6
CC FT CARBOHYD 210 210 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 677 677 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 803 AA; 90702 MW; D3A54E32958AED9 CRC64;

Query Match 10.3%; Score 156; DB 1; Length 803;
Best Local Similarity 24.6%; Pred. No. 0.00048;
Matches 58; Conservative 24; Mismatches 91; Indels 52; Gaps 12;

Oy 38 VSQCGCGCATCSDYNGCISCKSPKRFLPFALERIIGMKGQIGVCSSCPS--GYGTRYPDINK 95
Db 544 VNEECPSSCIVTEWEEWEECS---ATCRMQMKKGRMIMKMPADGSMCKADTEVEK 598
Oy 96 C--TRCK-----ADCDTGFNFKFCFKGSGFVHLHGKCLDNCPPEGLANN-- 138
Db 599 CMMPECHTIPCVLSPRSEWSDCSTGKGTRTRQR--MLKSPSELDGNEELBLKOWE 655
Oy 139 -HTWECVSVIHDCEVSNWPNPSPCTKKGKTCFGKRGTRTVEIIOHPSAKGNICPPNT 197
Db 656 CMLPCCP--ISCLTILWVSEC--NNSCG--KGMLRTRMTMEPORGAGCPETW 708
Oy 198 RKCVTORKKQKGERKGKGERKGKPKNGESKEAIPDSKLESKEPEORENKO 253
Db 709 KKCRL--RKCK-----SSGNERRHLK-----DAREKRSSEKIKEDSDGEO 747

RESULT 10
NTC2_MOUSE

ID NTC2_MOUSE STANDARD; PRT; 2470 AA.
 AC 035516; 060941; Q06008;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Neurogenic locus notch homolog protein 2 precursor (Notch 2) (Notch DE B1).
 GN NOTCH2.
 OS Mus musculus (Mouse).
 OC Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBi-TaxID=10090;
 RN [1]—
 RP SEQUENCE FROM N.A.
 RP STRAIN=CE7BL/6, TISSUE=Thymus.
 RA Hamada Y., Higuchi M., Tsujimoto Y.;
 RT "Complete amino acid sequence and multiform transcripts encoded by a single copy of mouse Notch2 gene."
 RT Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]—
 RP SEQUENCE OF 316-1518 FROM N.A.
 RP STRAIN=CE7BL/6 X CBA; TISSUE=embryo;
 RX MEDLINE=93178563; PubMed=8440332;
 RA Lardelli M., Lendahl U.;
 RT "Notch A and Notch B two mouse Notch homologues coexpressed in a wide variety of tissues,"
 RL Exp. Cell Res. 204:364-372(1993).
 RN [3]—
 RP SEQUENCE OF 1765-2153 FROM N.A.
 RP MEDLINE=9707510; PubMed=8917536;
 RA Milner L.A., Bigas R., Kopan R., Brasheen-Stein C., Bernstein I.D., Martin D.;
 RT "Inhibition of granulocytic differentiation by mNotch1.,"
 RL Proc. Natl. Acad. Sci. U.S.A. 93:13014-13019(1996).
 RN [4]—
 RP FUNCTION.
 RX MEDLINE=9396706; PubMed=10393120;
 RA Hamada Y., Kadokawa Y., Okabe M., Ikawa M., Coleman J.R., Tsujimoto Y.;
 RT "Mutation in ankyrin repeats of the mouse Notch2 gene induces early embryonic lethality,"
 RL Development 126:3415-3424(1995).
 RN [5]—
 RP DEVELOPMENTAL STAGE, AND ALTERNATIVE SPLICING.
 MEDLINE=95331893; PubMed=7609014;
 RA Higuchi M., Kiyama H., Hayakawa T., Hamada Y., Tsujimoto Y.;
 RT "Differential expression of Notch1 and Notch2 in developing and adult mouse brain,"
 RL Brain Res. Mol. Brain Res. 29:263-272(1995).
 RN [6]—
 RP POST-TRANSLATIONAL PROCESSING, AND MUTAGENESIS OF NET-1699.
 MEDLINE=21522956; PubMed=1511718;
 RA Saxena M., Schroeter E.H., Mumford J.S., Kopan R.;
 RT "Murine Notch homologs (N1-4) undergo presenilin-dependent proteolysis.,"
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9026-9031(2001).
 RL [7] Biol. Chem. 276:40268-0273(2001).
 RP POST TRANSLATIONAL PROCESSING, AND MUTAGENESIS OF NET-1699.
 MEDLINE=21237476; PubMed=11453941;
 RA Mizutani T., Taniguchi Y., Aoki T., Hashimoto N., Honjo T.;
 RT "Conservation of the biochemical mechanisms of signal transduction among mammalian Notch family members.,"
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9026-9031(2001).
 CC [1]— FUNCTION: Functions as a receptor for membrane-bound ligands Jagged1, Jagged2 and Delta1 to regulate cell-fate determination. Upon ligand activation through the released notch intracellular domain (NTCD) it forms a transcriptional activator complex with RBP-J kappa and activates genes of the enhancer of split locus. Affects the implementation of differentiation, proliferation and apoptotic programs (By similarity). May play an essential role in post-implantation development, probably in some aspect of cell specification and/or differentiation. CC [2]— SUBUNIT: Heterodimer of a C-terminal fragment N(TM) and a N-terminus of the protein. CC [3]— SUBUNIT: Heterodimer of a C-terminal fragment N(TM) and a N-terminus of the protein.

CC terminal fragment N(EC) which are probably linked by disulfide bonds.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Following proteolytic processing NTCD is translocated to the nucleus.
 CC -!- ALTERNATIVE PRODCS: 2 isoforms; 1 (shown here) and 2; may be produced by alternative splicing.
 CC -!- TISSUE SPECIFICITY: Expressed in the brain, liver, kidney, neuroepithelia, somites, optic vesicles and branchial arches, but not heart.
 CC -!- DEVELOPMENTAL STAGE: Expressed in the embryonic ventricular zone, the postnatal epndymal cells, and the choroid plexus throughout embryonic and postnatal development.
 CC -!- PTM: Synthesized in the endoplasmic reticulum as an inactive form which is proteolytically cleaved by a furin-like convertase in the trans-Golgi network before it reaches the plasma membrane to yield an active, ligand-accessible form. Cleavage results in C-terminal fragment N(TM) and a N-terminal fragment N(EC). Following ligand binding, it is cleaved by TNF-alpha converting enzyme (TACE) to yield a membrane-associated intermediate fragment called notch extracellular truncation (NET). This fragment is then cleaved by presenilin dependent gamma-secretase to release a notch-derived peptide containing the intracellular domain (NTCD) from the membrane.
 CC -!- PTM: Phosphorylated.
 CC -!- SIMILARITY: CONTAINS 34.5 EGFLIKE DOMAINS.
 CC -!- SIMILARITY: CONTAINS 6 ANK REPEATS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstations of the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC DR D32210; BAA22054.1;
 CC EMBL: X61279; CA4810.1;
 CC EMBL: U31881; AA52924.1; -.
 CC DR HSSP; P16109; 1FB2.
 CC DR MGI: 97364; Notch2.
 CC DR InterPro; IPR002110; ANK.
 CC DR InterPro; IPR005152; ASX_Hydroxyl.
 CC DR InterPro; IPR00561; EGFLike.
 CC DR InterPro; IPR00742; EGF_2.
 CC DR InterPro; IPR01881; EGF_Ca.
 CC DR InterPro; IPR00138; EGF_11.
 CC DR InterPro; IPR00080; Notch.
 CC DR Pfam; PF00008; EGF_35.
 CC DR Pfam; PF00023; ank; 6.
 CC DR Pfam; PF00066; notch; 2.
 CC DR PRINTS; PR00010; EGFLBLOOD.
 CC DR PRINTS; PR01452; NOTCH.
 CC DR SMART; SM00248; ANK; 4.
 CC DR SMART; SM00179; EGF_Ca; 22.
 CC DR SMART; SM00001; EGF_Like; 12.
 CC DR SMART; SM00094; NL_3.
 CC DR PROSITE; PS00088; ANK_REPEAT; 4.
 CC DR PROSITE; PS50297; ANK REP REGION; 1.
 CC DR PROSITE; PS00010; ASX_HYDROXYL; 22.
 CC DR PROSITE; PS0022; EGF_1; 33.
 CC DR PROSITE; PS01186; EGF_2; 27.
 CC DR PROSITE; PS01187; EGF_Ca; 22.
 CC DR Receptor; Transcription regulation; Activator; Differentiation; Development; Protein; Repeat; ANK domain; Transmembrane; Glycoprotein; Signal; Phosphorylation; Alternative splicing.
 CC FT SIGNAL 1 25 POTENTIAL.
 CC FT CHAIN 26 2470 NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 2.
 CC FT CHAIN 1666 2470 NOTCH EXTRACELLULAR TRUNCATION.
 CC FT DOMAIN 1697 2470 NOTCH INTRACELLULAR DOMAIN.
 CC FT DOMAIN 1677 26 EXTRACELLULAR (POTENTIAL).

PT	TRANSMEM	1678	1698	POTENTIAL.
FT	DOMAIN	1699	2470	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	182	63	EGF-LIKE 1.
FT	DOMAIN	64	102	EGF-LIKE 2.
FT	DOMAIN	105	143	EGF-LIKE 3.
FT	DOMAIN	144	180	EGF-LIKE 4.
FT	DOMAIN	182	219	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	221	256	EGF-LIKE 5 (INCOMPLETE).
FT	DOMAIN	258	294	EGF-LIKE 6 (INCOMPLETE).
FT	DOMAIN	295	334	EGF-LIKE 7.
FT	DOMAIN	335	372	EGF-LIKE 8. CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	373	411	EGF-LIKE 9. CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	413	452	EGF-LIKE 10.
FT	DOMAIN	454	490	EGF-LIKE 11. CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	492	528	EGF-LIKE 12. CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	520	566	EGF-LIKE 13. CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	566	603	EGF-LIKE 14. CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	605	641	EGF-LIKE 15. CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	643	678	EGF-LIKE 16. CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	676	716	EGF-LIKE 17. CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	718	753	EGF-LIKE 18. CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	755	791	EGF-LIKE 19. CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	793	829	EGF-LIKE 20. CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	831	869	EGF-LIKE 21. CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	871	907	EGF-LIKE 22.
FT	DOMAIN	909	945	EGF-LIKE 23. CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	947	983	EGF-LIKE 24. CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	985	1021	EGF-LIKE 25. CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1023	1059	EGF-LIKE 26. CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1061	1097	EGF-LIKE 27. CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1099	1145	EGF-LIKE 28.
FT	DOMAIN	1147	1183	EGF-LIKE 29. CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1185	1221	EGF-LIKE 30. CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1223	1260	EGF-LIKE 31. CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1262	1300	EGF-LIKE 32. CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1302	1345	EGF-LIKE 33.
FT	DOMAIN	1345	1410	EGF-LIKE 34.
FT	DOMAIN	1418	1454	EGF-LIKE 35.
FT	REPEAT	1503	1533	LIN/NOTCH 1.
FT	REPEAT	1825	1869	LIN/NOTCH 2.
FT	REPEAT	1874	1903	ANK 1.
FT	REPEAT	1907	1937	ANK 2.
FT	REPEAT	1941	1970	ANK 3.
FT	REPEAT	1974	2003	ANK 4.
FT	REPEAT	2007	2036	ANK 5.
FT	DOMAIN	1645	1648	POLY-ALA.
FT	DOMAIN	1992	1995	POLY-LEU.
FT	DOMAIN	2183	2189	POLY-ALA.
FT	DOMAIN	2185	2428	POLY-SER.
FT	DOMAIN	2445	2450	POLY-GLY.
FT	DISUFID	28	41	BY SIMILARITY.
FT	DISUFID	35	51	BY SIMILARITY.
FT	DISUFID	53	62	BY SIMILARITY.
FT	DISUFID	68	79	BY SIMILARITY.
Query	Match	9.9%	Score 150.5; DB 1; Length 2470;	
Qy	Best Local Similarity	24.3%	Pred No. 0.0031; Mismatches 53; Conservative 23; Indels 89; Gaps 12;	
Qy	Matches	53		
Qy	31	ORRMPNPSSQCGCAGCACSDYNGCCLKPFFALERIGMKGQIGVCISSCPGYYGTRY 90		
Db	945	QTDNECLSEPCNG-GCSDFVNSYC-----TCAGFHVHC 982		
Db	91	P-DINCKTCKKAKCDTCNKNFC-----TKCKSGFYUHLGK 125		
Db	983	ENNIDECTE-----SSCFNGTCVDGINSFSSLCLCPVGFITGPFLHDINECSSNPCLNAGT 1037		
Qy	126	CLDN-----CPGEGLEANNHMEVCSIVH-CESEENWPSPTCKKGKTCGKRGKETRV 177		
Db	1038	CVDGEGTYRCICPLGYTKN-----COTLNICS-----RSPCKNG-TC----- 1076		
Qy	178	REIJOHPSAKGNICPPNETKCTVORKKC-----QKG 210		
Db	1077	--VQEKARHCLCPPGWDGAYCDVJNVSCKAAQKG 1111		

DR Pfam; PP00069; Phosphatase; 1.
 DR Pfam; PF00757; Furin-like; 1.
 DR Pfam; PF00130; Recep L domain; 2.
 DR PRINTS; PRO0109; TYRKINASE.
 DR PRODOM; PD000001; Euk_pf kinase; 1.
 DR SMART; SMM0060; FN1; 3.
 DR SMART; SMM0219; TWRK; 1.
 DR PROSITE; PS00107; PROTEIN KINASE_ATP; 1.
 DR PROSITE; PS00109; PROTEIN KINASE_Tyr; 1.
 DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 KW Transferase; Tyrosine-protein kinase; Receptor; Transmembrane;
 KW Glycoprotein; ATP-binding; Phosphorylation; Repeat; Signal.
 FT SIGNAL 1 30 POTENTIAL.
 FT CHAIN 31 741 INSULIN-LIKE GROWTH FACTOR_I RECEPTOR,
 FT DOMAIN 742 1370 BETA-CHAIN.
 FT DOMAIN 742 936 INSULIN-LIKE GROWTH FACTOR_I RECEPTOR,
 FT DOMAIN 937 960 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 961 1370 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 609 929 FIBRONECTIN TYPE-III 1.
 FT DOMAIN 830 929 FIBRONECTIN TYPE-III 2.
 FT DOMAIN 1000 1275 PROTEIN KINASE.
 FT DOMAIN 1005 1014 ATP (BY SIMILARITY).
 FT BINDING 1034 1034 ATP (BY SIMILARITY).
 FT ACT_SITE 1136 1136 BY SIMILARITY.
 FT DISULFID 215 224 BY SIMILARITY.
 FT DISULFID 219 230 BY SIMILARITY.
 FT DISULFID 231 239 BY SIMILARITY.
 FT DISULFID 235 248 BY SIMILARITY.
 FT DISULFID 251 260 BY SIMILARITY.
 FT DISULFID 264 276 BY SIMILARITY.
 FT DISULFID 282 303 BY SIMILARITY.
 FT DISULFID 307 321 BY SIMILARITY.
 FT DISULFID 324 328 BY SIMILARITY.
 FT CARBOHYD 51 51 N-LINKED (GLCNAC . . .) (POTENTIAL).
 FT CARBOHYD 102 102 N-LINKED (GLCNAC . . .) (POTENTIAL).
 FT CARBOHYD 135 135 N-LINKED (GLCNAC . . .) (POTENTIAL).
 FT CARBOHYD 245 245 N-LINKED (GLCNAC . . .) (POTENTIAL).
 FT CARBOHYD 314 314 N-LINKED (GLCNAC . . .) (POTENTIAL).
 FT CARBOHYD 418 418 N-LINKED (GLCNAC . . .) (POTENTIAL).
 FT CARBOHYD 439 439 N-LINKED (GLCNAC . . .) (POTENTIAL).
 FT CARBOHYD 535 535 N-LINKED (GLCNAC . . .) (POTENTIAL).
 FT CARBOHYD 608 608 N-LINKED (GLCNAC . . .) (POTENTIAL).
 FT CARBOHYD 623 623 N-LINKED (GLCNAC . . .) (POTENTIAL).
 FT CARBOHYD 641 641 N-LINKED (GLCNAC . . .) (POTENTIAL).
 FT CARBOHYD 748 748 N-LINKED (GLCNAC . . .) (POTENTIAL).
 FT CARBOHYD 757 757 N-LINKED (GLCNAC . . .) (POTENTIAL).
 FT CARBOHYD 765 765 N-LINKED (GLCNAC . . .) (POTENTIAL).
 FT CARBOHYD 901 901 N-LINKED (GLCNAC . . .) (POTENTIAL).
 FT CARBOHYD 914 914 N-LINKED (GLCNAC . . .) (POTENTIAL).
 FT MOD_RSS 1166 1166 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CONFLICT 985 986 AD -> PY (IN REF. 3).
 SQ SEQUENCE 1370 AA: 155395 MW: 4594689741CB145 CRC64:

Query Match 9.8%; Score 19; DB 1; Length 1370;
 Best Local Similarity 26.7%; Pred. No. 0.0024;
 Matches 47; Conservative 17; Mismatches 46; Indels 66; Gaps 11;

QY 35 HPNVNSGCGCATCSDYNGCLSCPKRLPFFALERIGMKQIGVQLSPSGVY--GTRVP 91
 DB 232 HPE--CIGSCHPPDNTTCVACRIVYY-----KGFYLUHGKCLDNPEGLEANNHTM 143
 QY 92 DINKCTKCKRADCTCFNKGKFTKC-----KSGFYLHGKCLDNPEGLEANNHTM 143
 DB 276 -----CVDRDFCAMPINPNAESSSDGFRVHDGEWQECPSGFRNSTSOMYC 321
 QY 144 VSVIHCVESEWNPSPCTKGKGTGKGKGTETRIEIOPSA-----KGNL 190
 DB 322 IP--CE-----GPCP---KVCG--DEEKITKTTDSVTSQMQGCTILKGNL 361

RESULT 12
 NOTC_XENLA STANDARD; PRT; 2524 AA.
 ID NOTC_XENLA STANDARD; PRT; 2524 AA.
 AC P21783;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Neurogenic locus notch protein homolog precursor (NOTCH protein).
 GN XOTCH.
 OS Xenopus levis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Messobatrachia; Pipidae;
 Xenopodidae; Xenopus.
 OC NOBI_TaxID=8355;
 RN [1] SEQUENCE FROM N.A.
 RP MEDLINE=90385285; PubMed=2402639;
 RA Coffman C, Harris W, Kintner C;
 RL "Notch, the Xenopus homolog of Drosophila notch. ";
 RL Science 249:1438-1441(1990).
 RN [2] REVISONS TO 1759-1782.
 RA Kintner C.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 CC --!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC --!- DEVELOPMENTAL STAGE: EXPRESSED ALMOST UNIFORMLY IN EARLY EMBRYOS.
 CC --!- SIMILARITY: BELONGS TO THE NOTCH FAMILY.
 CC --!- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
 CC --!- SIMILARITY: CONTAINS 3 LIN-NOTCH REPEATS.
 CC --!- SIMILARITY: CONTAINS 6 ANK REPEATS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC or send an email to license@isb-sib.ch).
 DR EMBL; M30374; AAB02039.1; -.
 DR PIR; A35844; A35844
 DR HSSP; P00740; IEDM.
 DR Interpro; IPR002110; ANK.
 DR Interpro; IPR00152; Asx hydroxyl.
 DR Interpro; IPR00561; EGF-like.
 DR Interpro; IPR00742; EGF-2.
 DR Interpro; IPR01881; EGF_Ca.
 DR Interpro; IPR01438; EGF_II.
 DR Interpro; IPR00800; NOTCH.
 DR Pfam; PF00008; EGF; 36.
 DR Pfam; PF00023; ank; 6.
 DR Pfam; PF00066; notch; 3.
 DR PRINTS; PR00010; EGBLOOD.
 DR PRINTS; PR00142; NOTCH.
 DR SMART; SMM0248; ANK; 5.
 DR SMART; SMM0119; EGF_Ca; 23.
 DR SMART; SMM0001; EGF-like; 11.
 DR SMART; SMM0004; NL; 2.
 DR PROSITE; PS50088; ANK_REPEAT; 4.
 DR PROSITE; PS50297; ANK_REPEAT; 1.
 DR PROSITE; PS00110; ASX_HYDROXYL; 23.
 DR PROSITE; PS00022; EGF_1; 34.
 DR PROSITE; PS01186; EGF_2; 29.
 DR PROSITE; PS01187; EGF_Ca; 21.
 DR Differentiation; Neurogenesis; Repeat; ANK repeat; EGF-like domain;
 KW Transmembrane; Signal; Glycoprotein.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 2524 NEUROGENIC LOCUS NOTCH PROTEIN HOMOLOG.
 FT DOMAIN 20 1728 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1729 1750 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 1751 2524 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 1752 57 EGF-LIKE 1.

FT	DOMAIN	58	99	EGF-LIKE 2.	FT	DISULFID	399	408	BY SIMILARITY.
FT	DOMAIN	102	140	EGF-LIKE 3.	FT	DISULFID	415	428	BY SIMILARITY.
FT	DOMAIN	141	177	EGF-LIKE 4.	FT	DISULFID	422	437	BY SIMILARITY.
FT	DOMAIN	179	215	EGF-LIKE 5.	FT	DISULFID	439	448	BY SIMILARITY.
FT	DOMAIN	217	254	CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	455	466	BY SIMILARITY.
FT	DOMAIN	256	292	EGF-LIKE 6.	FT	DISULFID	460	475	BY SIMILARITY.
FT	DOMAIN	294	332	EGF-LIKE 7.	FT	DISULFID	477	485	BY SIMILARITY.
FT	DOMAIN	334	370	CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	493	504	BY SIMILARITY.
FT	DOMAIN	371	409	EGF-LIKE 8.	FT	DISULFID	498	513	BY SIMILARITY.
FT	DOMAIN	411	449	EGF-LIKE 9.	FT	DISULFID	515	524	BY SIMILARITY.
FT	DOMAIN	451	487	EGF-LIKE 10.	FT	DISULFID	531	542	BY SIMILARITY.
FT	DOMAIN	489	525	EGF-LIKE 11.	FT	DISULFID	536	551	BY SIMILARITY.
FT	DOMAIN	527	563	EGF-LIKE 12.	FT	DISULFID	553	562	BY SIMILARITY.
FT	DOMAIN	565	600	EGF-LIKE 13.	FT	DISULFID	562	579	BY SIMILARITY.
FT	DOMAIN	602	638	EGF-LIKE 14.	FT	DISULFID	574	588	BY SIMILARITY.
FT	DOMAIN	640	675	EGF-LIKE 15.	FT	DISULFID	590	599	BY SIMILARITY.
FT	DOMAIN	673	713	EGF-LIKE 16.	FT	DISULFID	606	617	BY SIMILARITY.
FT	DOMAIN	715	750	EGF-LIKE 17.	FT	DISULFID	611	626	BY SIMILARITY.
FT	DOMAIN	752	788	EGF-LIKE 18.	FT	DISULFID	628	637	BY SIMILARITY.
FT	DOMAIN	790	826	EGF-LIKE 19.	FT	DISULFID	644	654	BY SIMILARITY.
FT	DOMAIN	828	866	EGF-LIKE 20.	FT	DISULFID	649	663	BY SIMILARITY.
FT	DOMAIN	866	906	EGF-LIKE 21.	FT	DISULFID	665	674	BY SIMILARITY.
FT	DOMAIN	906	942	EGF-LIKE 22.	FT	DISULFID	681	692	BY SIMILARITY.
FT	DOMAIN	944	980	EGF-LIKE 23.	FT	DISULFID	686	701	BY SIMILARITY.
FT	DOMAIN	982	1018	EGF-LIKE 24.	FT	DISULFID	703	712	BY SIMILARITY.
FT	DOMAIN	1020	1055	EGF-LIKE 25.	FT	DISULFID	719	729	BY SIMILARITY.
FT	DOMAIN	1058	1094	EGF-LIKE 26.	FT	DISULFID	724	738	BY SIMILARITY.
FT	DOMAIN	1096	1142	EGF-LIKE 27.	FT	DISULFID	740	749	BY SIMILARITY.
FT	DOMAIN	1144	1180	EGF-LIKE 28.	FT	DISULFID	756	767	BY SIMILARITY.
FT	DOMAIN	1182	1218	EGF-LIKE 29.	FT	DISULFID	761	776	BY SIMILARITY.
FT	DOMAIN	1220	1264	EGF-LIKE 30.	FT	DISULFID	778	787	BY SIMILARITY.
FT	DOMAIN	1266	1304	EGF-LIKE 31.	FT	DISULFID	794	805	BY SIMILARITY.
FT	DOMAIN	1306	1346	EGF-LIKE 32.	FT	DISULFID	805	814	BY SIMILARITY.
FT	DOMAIN	1347	1383	EGF-LIKE 33.	FT	DISULFID	816	825	BY SIMILARITY.
FT	DOMAIN	1386	1424	EGF-LIKE 34.	FT	DISULFID	832	843	BY SIMILARITY.
FT	REPEAT	1441	1478	LIN/NOTCH 1.	FT	DISULFID	856	865	BY SIMILARITY.
FT	REPEAT	1479	1520	LIN/NOTCH 2.	FT	DISULFID	872	883	BY SIMILARITY.
FT	REPEAT	1521	1560	LIN/NOTCH 3.	FT	DISULFID	877	892	BY SIMILARITY.
FT	REPEAT	1876	1919	ANK 1.	FT	DISULFID	894	903	BY SIMILARITY.
FT	REPEAT	1924	1953	ANK 2.	FT	DISULFID	921	931	BY SIMILARITY.
FT	REPEAT	1957	1987	ANK 3.	FT	DISULFID	930	941	BY SIMILARITY.
FT	REPEAT	1991	2020	ANK 4.	FT	DISULFID	941	951	BY SIMILARITY.
FT	REPEAT	2024	2053	ANK 5.	FT	DISULFID	952	962	BY SIMILARITY.
FT	REPEAT	2057	2086	ANK 6.	FT	DISULFID	967	977	BY SIMILARITY.
FT	DISULFID	22	35	BY SIMILARITY.	FT	DISULFID	991	1006	BY SIMILARITY.
FT	DISULFID	29	45	BY SIMILARITY.	FT	DISULFID	1008	1017	BY SIMILARITY.
FT	DISULFID	47	56	BY SIMILARITY.	FT	DISULFID	1024	1035	BY SIMILARITY.
FT	DISULFID	62	74	BY SIMILARITY.	FT	DISULFID	1029	1044	BY SIMILARITY.
FT	DISULFID	68	87	BY SIMILARITY.	FT	DISULFID	1046	1055	BY SIMILARITY.
FT	DISULFID	89	98	BY SIMILARITY.	FT	DISULFID	1062	1073	BY SIMILARITY.
FT	DISULFID	106	117	BY SIMILARITY.	FT	DISULFID	1067	1082	BY SIMILARITY.
FT	DISULFID	111	128	BY SIMILARITY.	FT	DISULFID	1084	1093	BY SIMILARITY.
FT	DISULFID	130	139	BY SIMILARITY.	FT	DISULFID	1100	1121	BY SIMILARITY.
FT	DISULFID	145	156	BY SIMILARITY.	FT	DISULFID	1115	1130	BY SIMILARITY.
FT	DISULFID	150	165	BY SIMILARITY.	FT	DISULFID	1132	1141	BY SIMILARITY.
FT	DISULFID	167	175	BY SIMILARITY.	FT	DISULFID	1148	1159	BY SIMILARITY.
FT	DISULFID	183	194	BY SIMILARITY.	FT	DISULFID	1153	1168	BY SIMILARITY.
Query Match 9.8%; Score 148.5; DB 1; Length 2524; Best Local Similarity 22.5%; Pid: No. 0.0043; Matches 73; Conservative 32; Mismatches 102; Indels 117; Gaps 19; Ov 34 MHPVNSQCGCCTGDSYNGCLSKPRLFFALERGM-----KQGV 76									
Db	869	MNECNVRCPNG-ATQONTNGSYKNCRKPGYFGRNEMDIDCOPNPCKNGGSDGIM	927						
Ov	77	CLSSCPSCYYGTR-YPDINCKTK--CK--ADDDTCNKNCTKKGFGV-LH-----	123						
Db	928	FFCNCAGPFRGPCKEEDINECASNPCKNGANCTDVCNSYCT-CQPGSGIHCESNTPPDC	986						
Ov	124	--GKCLD-----NCPEGLAN--NHTTM-ECVSI-----	146						
Db	987	TESSCFNGGTCIDGINTFTCQCPGFGTSCYCOHDINECDSKPCLNGTCQDSYGTYKCTC	1046						

PTM: phosphorylated (By similarity).
SIMILARITY: BELONGS TO THE NOTCH FAMILY.
SIMILARITY: CONTAINS 35 EGF-LIKE DOMAINS.
SIMILARITY: CONTAINS 2 LIN-NOTCH REPEATS.
SIMILARITY: CONTAINS 6 ANK REPEATS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>). Send an email to licenses@isb-sib.ch.

FT	DOMAIN	795	831	EGF-LIKE 21, CALCIUM-BINDING (POTENTIAL).	Matches	54	conservative	23	Mismatches	52	Indels	89	Gaps	13;
FT	DOMAIN	823	871	EGF-LIKE 23, CALCIUM-BINDING (POTENTIAL).	OY	31	QRRMMPNYSQCGCGCATCSYNGCLCKPRLFPLALERIGMKQIGVCLSSCPGSGYGRY	90						
FT	DOMAIN	873	909	EGF-LIKE 24, CALCIUM-BINDING (POTENTIAL).	OY	1	;	;	;	;	;	;	;	
FT	DOMAIN	911	947	EGF-LIKE 24, CALCIUM-BINDING (POTENTIAL).	Db	947	OTDNECLSEPCPGAG-GTCSDVYNSYTC							-TCPAGFHGHC 984
FT	DOMAIN	949	985	EGF-LIKE 25, CALCIUM-BINDING (POTENTIAL).	QY	91	P-DIKTCICKADCPFNKQFC-----TKQSGFYLHCK 125							
FT	DOMAIN	987	1023	EGF-LIKE 26, CALCIUM-BINDING (POTENTIAL).	QY		;	;	;	;	;	;		
FT	DOMAIN	1025	1061	EGF-LIKE 27, CALCIUM-BINDING (POTENTIAL).	QY		;	;	;	;	;	;		
FT	DOMAIN	1063	1099	EGF-LIKE 28.	Db	985	ENNIDCETE-----SSFPNGTCVQGINSFSCLPQVSGFTGPFCPLHDINECSSNPCLNSGT	1039						
FT	DOMAIN	1101	1147	EGF-LIKE 29.	OY		;	;	;	;	;	;		
FT	DOMAIN	1149	1185	EGF-LIKE 30, CALCIUM-BINDING (POTENTIAL).	OY	126	CLD-----NCPEGLEANNITMCEWSVTH-CEVSENPNWPCTKKGKTCGFKRGTFRV	177						
FT	DOMAIN	1187	1223	EGF-LIKE 32, CALCIUM-BINDING (POTENTIAL).	Db	1040	CVDGIGTVRCTCPGLYTGKNC-----CQTLVNL-----SP-SPCKNGK-TCA-----	1079						
FT	DOMAIN	1225	1262	EGF-LIKE 33.	OY		;	;	;	;	;	;		
FT	DOMAIN	1264	1302	EGF-LIKE 34.	Db	1080	-----OEKARPRCLCPGWDGAYCDVLYNVSKAAALQKG	1113						
FT	DOMAIN	1304	1343	EGF-LIKE 35.	OY	178	REIICHPASAKGNLCPNNETRKCTVORKKC-----QKG	210						
FT	DOMAIN	1343	1412	POLY-ALA.	Db		;	;	;	;	;	;		
FT	DOMAIN	1446	1456	POLY-LEU.	OY		;	;	;	;	;	;		
FT	REPEAT	1420	1456	POLY-SER.	Db		;	;	;	;	;	;		
FT	REPEAT	1456	1546	POLY-GLY.	OY		;	;	;	;	;	;		
FT	REPEAT	1503	1535	LIN/NOTCH 1.	Db		;	;	;	;	;	;		
FT	REPEAT	1535	1571	LIN/NOTCH 2.	OY		;	;	;	;	;	;		
FT	REPEAT	1827	1871	ANK 1.	Db		;	;	;	;	;	;		
FT	REPEAT	1876	1905	ANK 2.	OY		;	;	;	;	;	;		
FT	REPEAT	1909	1939	ANK 3.	Db		;	;	;	;	;	;		
FT	REPEAT	1943	1972	ANK 4.	OY		;	;	;	;	;	;		
FT	REPEAT	1976	2005	ANK 5.	Db		;	;	;	;	;	;		
FT	REPEAT	2009	2038	ANK 6.	OY		;	;	;	;	;	;		
FT	DISULFID	28	41	BY SIMILARITY.	Db		;	;	;	;	;	;		
FT	DISULFID	35	51	BY SIMILARITY.	OY		;	;	;	;	;	;		
FT	DISULFID	53	62	BY SIMILARITY.	Db		;	;	;	;	;	;		
FT	DISULFID	68	79	BY SIMILARITY.	OY		;	;	;	;	;	;		
FT	DISULFID	73	90	BY SIMILARITY.	Db		;	;	;	;	;	;		
FT	DISULFID	90	101	BY SIMILARITY.	OY		;	;	;	;	;	;		
FT	DISULFID	109	121	BY SIMILARITY.	Db		;	;	;	;	;	;		
FT	DISULFID	115	131	BY SIMILARITY.	OY		;	;	;	;	;	;		
FT	DISULFID	133	142	BY SIMILARITY.	Db		;	;	;	;	;	;		
FT	DISULFID	148	159	BY SIMILARITY.	OY		;	;	;	;	;	;		
FT	DISULFID	153	168	BY SIMILARITY.	Db		;	;	;	;	;	;		
FT	DISULFID	170	179	BY SIMILARITY.	OY		;	;	;	;	;	;		
FT	DISULFID	186	198	BY SIMILARITY.	Db		;	;	;	;	;	;		
FT	DISULFID	192	207	BY SIMILARITY.	OY		;	;	;	;	;	;		
FT	DISULFID	209	218	BY SIMILARITY.	Db		;	;	;	;	;	;		
FT	DISULFID	225	236	BY SIMILARITY.	OY		;	;	;	;	;	;		
FT	DISULFID	230	246	BY SIMILARITY.	Db		;	;	;	;	;	;		
FT	DISULFID	248	257	BY SIMILARITY.	OY		;	;	;	;	;	;		
FT	DISULFID	254	275	BY SIMILARITY.	Db		;	;	;	;	;	;		
FT	DISULFID	269	284	BY SIMILARITY.	OY		;	;	;	;	;	;		
FT	DISULFID	286	295	BY SIMILARITY.	Db		;	;	;	;	;	;		
FT	DISULFID	302	315	BY SIMILARITY.	OY		;	;	;	;	;	;		
FT	DISULFID	309	324	BY SIMILARITY.	Db		;	;	;	;	;	;		
FT	DISULFID	326	335	BY SIMILARITY.	OY		;	;	;	;	;	;		
FT	DISULFID	342	353	BY SIMILARITY.	Db		;	;	;	;	;	;		
FT	DISULFID	347	362	BY SIMILARITY.	OY		;	;	;	;	;	;		
FT	DISULFID	364	373	BY SIMILARITY.	Db		;	;	;	;	;	;		
FT	DISULFID	379	390	BY SIMILARITY.	OY		;	;	;	;	;	;		
FT	DISULFID	384	401	BY SIMILARITY.	Db		;	;	;	;	;	;		
FT	DISULFID	403	412	BY SIMILARITY.	OY		;	;	;	;	;	;		
FT	DISULFID	419	433	BY SIMILARITY.	Db		;	;	;	;	;	;		
FT	DISULFID	427	442	BY SIMILARITY.	OY		;	;	;	;	;	;		
FT	DISULFID	444	453	BY SIMILARITY.	Db		;	;	;	;	;	;		
FT	DISULFID	450	471	BY SIMILARITY.	OY		;	;	;	;	;	;		
FT	DISULFID	465	480	BY SIMILARITY.	Db		;	;	;	;	;	;		
FT	DISULFID	482	491	BY SIMILARITY.	OY		;	;	;	;	;	;		
FT	DISULFID	498	509	BY SIMILARITY.	Db		;	;	;	;	;	;		
FT	DISULFID	518	529	BY SIMILARITY.	OY		;	;	;	;	;	;		
FT	DISULFID	520	547	BY SIMILARITY.	Db		;	;	;	;	;	;		
FT	DISULFID	536	556	BY SIMILARITY.	OY		;	;	;	;	;	;		
FT	DISULFID	558	567	BY SIMILARITY.	Db		;	;	;	;	;	;		
FT	DISULFID	574	584	BY SIMILARITY.	OY		;	;	;	;	;	;		

SEQUENCE FROM N.A. (ISOFORMS A; B; C AND D), AND FUNCTION.

STRAIN=Bristol N2;

MEDLINE=95293228; Pubmed=1774813;

RA Thacker C., Peters K.W., Stravko M., Rose A.M.;

"The bli-4 locus of *Caenorhabditis elegans* encodes structurally distinct kex2/autolin-1-like endopeptidases essential for early development and adult morphology.";

RT Genes Dev. 9:956-971(1995).

[2]

SEQUENCE FROM N.A., AND ALTERNATIVE SPlicing.

RC STRAIN=Bristol N2;

RA Leterrier P., Mansley P.;

RL Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.

CC !- FUNCTION: Isoform A is required for normal production or maintenance of the adult cuticle. There is functional redundancy between the isoforms. Vital for embryonic and larval development.

CC !- TISSUE SPECIFICITY: In larvae and adults, expressed in all hypodermal cells, vulva and ventral nerve cords.

CC !- DEVELOPMENTAL STAGE: Expression starts at embryo two-fold stage through to adults.

CC !- ALTERNATIVE PRODUCTS: 5 isoforms; A/a, B/b, C/c, D/d (shown here) and e; may be produced by alternative splicing.

CC !- SIMILARITY: BELONGS TO PEPTIDE FAMILY SB, FURIN SUBFAMILY.

CC !- CAUTION: Ref.1 sequence differs from that shown due to a frameshift in position 631. CC shown due to erroneous gene model prediction.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

Query Match Score 1.7%; DB 1; Length 2471;

Best Local Similarity 24.8%; Pred. No. 0.005;

FT	CARBONYD	88	89	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBONID	245	245	N-LINKED (GLCNAC. . .)	(POTENTIAL).
SO	SBROUKE	379	AA:	41590	MW: E3765F5642BCCA CRC64;
QY	Query	Match	9.5%	Score	143.5; DB 1; Length 379;
QY	Best	Local Similarity	24.3%	Score	143.5; DB 1; Length 379;
QY	Matches	60;	Conservative	24;	Mismatches 68; Indels 95; Caps 17;
QY	42	CQGGAT--CSDYNGC-----LSCKPRLLFALERIAGMKQIQGVCHSS---C	81		
Db	182	CPGGCRNGFCBERRVCBPGFQPHCERKALCIPRCMG-----	182	: :	: :
Qy	82	PSGYVTRYPLDINKTCR-KDCCD-TCFVKRNTRTKCSGFYLHIGKCLNDNCBPEGLEANNH	139	: :	: :
Db	233	PPGTYG-----VNCDAKNCSTCFNGGTC-----FV--PERCI--CPGPLEEE-	271	: :	: :
Qy	140	TMECVSIVHCEYSEWNPMSPCTKKGKTCFKRTRTTRVREIIQHPSA-KGNNLCP-----	192	: :	: :
Db	272	-----QCELSKCP--QPCRNGGKCIG-----KSKKCPCPKGYODLDSKPVCEP	312	: :	: :
Qy	193	-----PTNETRKCCTYORKKCKQKGERGKGGRER-----KKPKMGKESKEAIPDSKLE	240	: :	: :
Db	313	GCGAHCCTGTCBPNKC-----QCREGWHGRHCKNRYGASLMHAPRPGAGLERHTPSLKKA	367	: :	: :
Qy	241	SSKIKE	247	:	:
Db	368	DRRDPPE	374	:	:

Search completed: May 6, 2003, 14:52:00
Job time : 16 secs

GenCore version 5.1.4.p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2003, 14:50:38 ; Search time 18 Seconds

(Without alignments) 1458.039 Million cell updates/sec

Title: US-09-894-912A-13

Perfect score: 1516 MGHIRLILWLFILNFEHYI.....QOKKRKVQDKQKSVSVSTVH 273

Scoring table: BLOSUM62 Gapop 10.0 , Gapext 0.5

283224 seqs, 96134422 residues

Searched: Total number of hits satisfying chosen parameters:

283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database : PIR 73; * 1: pir1; * 2: pir2; * 3: pir3; * 4: pir4; *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Query Match Length DB ID

Description

Result No.	Score	Query	Match	Length	DB	ID	Description
1	192	12.7	1299	2	TA43251		furin (EC 3.4.21.7)
2	186.5	12.3	1548	2	S34583		furin (EC 3.4.21.7)
3	184	12.3	962	2	JCS571		furin (EC 3.4.21.7)
4	184	12.1	975	2	JCS570		furin (EC 3.4.21.7)
5	176.5	11.6	1680	2	A43434		furin (EC 3.4.21.7)
6	173.5	11.4	969	1	A39490		furin (EC 3.4.21.7)
7	168	11.1	915	2	B48225		furin (EC 3.4.21.7)
8	167.5	11.0	932	2	T52525		furin (EC 3.4.21.7)
9	167	11.0	915	1	A48225		furin (EC 3.4.21.7)
10	162	10.7	899	2	GO2428		furin (EC 3.4.21.7)
11	162	10.7	915	2	JCG148		furin (EC 3.4.21.7)
12	161	10.6	440	2	T24222		furin (EC 3.4.21.7)
13	158.5	10.5	379	2	A59180		furin (EC 3.4.21.7)
14	156.5	10.3	937	2	A53282		furin (EC 3.4.21.7)
15	156	10.3	803	2	A47723		furin (EC 3.4.21.7)
16	153.5	10.1	942	2	D8703		furin (EC 3.4.21.7)
17	150.5	9.9	1203	2	A49175		furin (EC 3.4.21.7)
18	149	9.8	1371	2	A33837		furin (EC 3.4.21.7)
19	148.5	9.8	2524	2	A35944		furin (EC 3.4.21.7)
20	147.5	9.7	2471	2	A49128		furin (EC 3.4.21.7)
21	145	9.6	570	2	T37314		furin (EC 3.4.21.7)
22	145	9.6	1620	2	T27783		furin (EC 3.4.21.7)
23	143.5	9.5	807	2	A38152		furin (EC 3.4.21.7)
24	142	9.4	378	2	B59180		furin (EC 3.4.21.7)
25	142	9.4	677	2	C42125		furin (EC 3.4.21.7)
26	139	9.2	1367	1	IGHUR1		furin (EC 3.4.21.7)
27	139	9.2	2555	2	A40043		furin (EC 3.4.21.7)
28	138.5	9.1	1369	2	S70713		furin (EC 3.4.21.7)
29	138	9.1	2531	2	S18188		furin (EC 3.4.21.7)

ALIGNMENTS

RESULT 1

143251 Furin (EC 3.4.21.7) - fall armyworm

N;Alternate names: paired basic amino acid cleaving enzyme; proprotein convertase; seri

C;Species: Spodoptera frugiperda (fall armyworm)

C;Accession: T43251

R;Cipollik, M.; Klenk, H.

Submitted to the EMBL Data Library, January 1996

A;Description: Cloning and functional characterization of FURIN from Spodoptera frugiper

A;Reference number: 222368

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-1299 <CIE>

A;Cross-references: EMBL:Z68888; NID:91167859; PID:e219690; PMID:CAA93116.1

A;Experimental source: clone Sturin 6; ovary

C;Function:

A;Description: responsible for the endoproteolytic processing of proproteins with speci

C;Keywords: hydrolase; serine proteinase

Query Match Best Local Similarity 12.7%; Score 192; DB 2; Length 1299;

Matches 63; Conservative 26; Mismatches 78; Indels 60; Gaps 12;

Db 795 SVCRPCAAHATCSEERADGCTSCBHHL-----VLDGTCMASCPSPHYET--EDDM 843

Oy 96 CTGCKAQADDTF--NNPCTKCKCSGPGVHLGKCLDPCPELEANNNHBTBCVSVLHCVSE 153

Db 844 CAKCHESCDTQGPBTQCVTCRSTALDRGCVTSPPAYVADKKRKEC--MRCPVG- 899

Oy 154 WNPMSPTKKTCFPRKETRTRVREIQHSAKGNLCP--TNETKQC-TWORKKKQ 209

Db 900 ---CSTCH-----SAFCLSCPCPWELNKKGKCPMPGSDKSA 933

Oy 210 GENGKKGKERRERKPKK-----GESKE--AIPSKSLESSEKPE 247

Db 934 GEFAV--DQCKRCNPACDSGYGENGHCLTCPNPENLQDYKCP 977

Notch-1 protein - transmembrane protein
insulin receptor - insulin receptor
hypothetical protein
insulin receptor p
insulin receptor
insulin receptor
insulin receptor
insulin receptor - apoptosis-mediatin
insulin receptor
Notch homolog M0cc
hypothetical prote
trophozoite surface
furin (EC 3.4.21.7)
cysteine rich prot

A;Reference number: S34583; PMID:8335106

A:Accession:	S34583	Db	767	CVTLCPAGFYADE--SOKNLCKHPSKKVQDPEKETVCKEGFSLRASCIPECTY	824
A:Status:	preliminary	Qy	136	ANNHMEC-----VSI VHCEVS-EWNWNSPCTKKGKTCGPKRGTRVRE	179
A:Residue:	1-1548 <NAK>	Db	825	FDSELIRCGECHHTCGTCVGFGRRECCHGAKNHFHW---KCVPACGEGFPE---	875
A;Cross-references:	GB:D17583; NID:9407344; PIDN:BAA04507.1; PID:di005033; PID:9440374	C;Keywords:	hydrolase; serine proteinase		
Query Match	12.3%; Score 186.5; DB 2; Length 1548; Best Local Similarity 25.0%; Pred. No. 6.3e-05; Indels 90; Gaps 15; Matches 65; Conservative 31; Mismatches 90; Indels 77; Gaps 15;	Qy	39	SOGQGGCATCSDPY--CLSCPKPLFALERIGMKQIGVYSSCPGKVGYTRPDI--N	94
Db	880 NQPHSSCKTC--NGSLCASCPTCNYLWQ-----ACVPSCOG---TWPSVITSG	924			
Qy	95 KCTCKKAPEDCCTPNKNKTKKS---GFVTHLCKIDNCPEGLENNHME-CVSTIHC	149			
Db	925 SCKECKSCDVSCGGADLQQLCULSDFDNTLILHREGYHSCPGFYKDGVCHEHSS--PC	982			
Qy	150 EVSKNPNSPCTKKG-----KTC-----GKRGTRVREIOPASKNI- 190				
Db	983 KTCFEGNATCSNCSEGGDFVLDHGVCWKTCPERKHAVEGVCKHCPERQDCHEKTCCEMP	1042			
Qy	191 -----CPT--NERKCTVQRKCKQCKGGRKRGKGRKRPKNGESKAIRDS 236				
Db	1043 DFFFLYNDMCHRSPKSYDPMQHCPKHC-----LECGNPKEDDCKVCAATS	1091			
Qy	237 KSISS--KEIPE--QRENQ 253				
Db	1092 KALHNLCLDECPEGTYKEEENDE 1115				
RESULT 3					
subtilisin-like proprotein convertase (EC 3.4.21.-) PACE4 precursor, splice form E-II -					
C:Species:	Homo sapiens (man)	C:Species:	Homo sapiens (man)		
C:Date:	23-Sep-1997 #sequence_revision 23-Sep-1997 #text_change 20-Jun-2000	C:Date:	23-Sep-1997 #sequence_revision 23-Sep-1997 #text_change 20-Jun-2000		
C:Accession:	JCC5571	C:Accession:	JCC5570		
R:Mori, K.; Kii, S.; Tsuji, A.; Nagahama, M.; Iimamaki, A.; Hayashi, K.; Akanatsu, T.; Nag	R:Mori, K.; Kii, S.; Tsuji, A.; Nagahama, M.; Iimamaki, A.; Hayashi, K.; Akanatsu, T.; Nag	A:Title:	A novel human PACE4 isoform, PACE4E is an active processing protease containing		
J: Biochem. 131, 941-948, 1997	J: Biochem. 131, 941-948, 1997	A:Reference number:	JCC5570; MUID:97335942; PMID:9192737		
A:Accession:	JCC5571	A:Accession:	JCC5570		
A:Residue:	nucleic acid sequence not shown	A:Residue:	nucleic acid sequence not shown		
A:Gene:	GB:PACE4	A:Gene:	GB:PACE4		
A:Cross-references:	DBJ:087993; NID:92330548; PIDN:BAA21791.1; PID:92330549	A:Cross-references:	DBJ:087993; NID:92330548; PIDN:BAA21791.1; PID:92330549		
A:Experimental source:	brain cerebellum	A:Experimental source:	brain cerebellum		
C:Comment:	This enzyme is a processing protease and responsible for processing of various substrates intracellularly.	C:Comment:	This enzyme is a processing protease and responsible for processing of various substrates intracellularly.		
C:Genetics:		C:Genetics:			
A:Gene:	GB:PACE4	A:Gene:	GB:PACE4		
A:Cross-references:	GBD:131390; OMIM:167405	A:Cross-references:	GBD:131390; OMIM:167405		
A:Map position:	1q26.1-15q26	A:Map position:	1q26.1-15q26		
C:Superfamily:	subtilisin-like proteinase PACE4; subtilisin homology	C:Superfamily:	subtilisin-like proteinase PACE4; subtilisin homology		
C:Experimental source:	brain cerebellum	C:Experimental source:	brain cerebellum		
C:Comment:	This enzyme is a processing protease and responsible for processing of various substrates intracellularly.	C:Comment:	This enzyme is a processing protease and responsible for processing of various substrates intracellularly.		
A:Residue:	1-962 <MOR>	A:Residue:	1-962 <MOR>		
A:Map position:	1q26.1-15q26	A:Map position:	1q26.1-15q26		
C:Species:	subtilisin-like proteinase PACE4: subtilisin homology	C:Species:	subtilisin-like proteinase PACE4: subtilisin homology		
C:Keywords:	glycoprotein; hydrolase; serine proteinase	C:Keywords:	glycoprotein; hydrolase; serine proteinase		
F:1-62/Domain:	signal sequence	F:1-62/Domain:	signal sequence		
F:1-62/Domain:	#status predicted <SIG>	F:1-62/Domain:	#status predicted <SIG>		
F:63-149/Domain:	propeptide #status predicted <PRO>	F:63-149/Domain:	propeptide #status predicted <PRO>		
F:155-246/Domain:	hydrophobic cluster #status predicted <HCL>	F:155-246/Domain:	hydrophobic cluster #status predicted <HCL>		
F:259/Binding site:	carbohydrate (Asn) (covalent) #status predicted	F:259/Binding site:	carbohydrate (Asn) (covalent) #status predicted		
Query Match	12.1%; Score 184; DB 2; Length 975; Best Local Similarity 26.6%; Pred. No. 6.5e-05; Indels 62; Conservative 24; Mismatches 87; Indels 60; Gaps 14; Matches 62; Conservative 24; Mismatches 87; Indels 60; Gaps 14;	Qy	19 YIGSONASGRQRORRHMHPNVSGCGCATCSD--YNGCLSKPPLFALERIGMKQIGV	76	
Db	737 YFGDTAARRCR-----CHKGCTCSSRAATCQLCSR-RGYF---HHQEMNT	779			
Qy	77 CLSSPSGKVGYTRPDKNCKTCCKADDCDFCNK-NFTKCKSGFYLHLGKCLDCCPDELE	135			
Db	780 CVTLCPAGFYADE--SOKNLCKHPSKKVQDPEKETVCKEGFSLRASCIPECTY	837			
Qy	136 ANNHMEC-----VSI VHCEVS-EWNWNSPCTKKGKTCGPKRGTRVRE	179			
Db	838 FDSELIRCGECHHTCGTCVGFGRRECCHGAKNHFHW---KCVPACGEGFPE---	888			
Qy	180 1IQHPSAKNLCK-----PPTNTRKTVQKKCKGKGKRRERKPKNGE	228			
Db	889 --EMGPLPHKVCRRYSPPGGE-ROAVTS---SKGVPG--GOSLSASPGAE	932			
RESULT 5					
A:Accession:	A3434	A:Accession:	A3434		
C:Species:	Drosophila melanogaster	C:Species:	Drosophila melanogaster		
C:Date:	04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 31-Mar-2000	C:Date:	04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 31-Mar-2000		
C:Accession:	A3434	C:Accession:	A3434		
R:Roobroek, A.J.; Creemers, J.W.; Pauli, I.G.; Kurzik-Dumke, U.; Rentrop, M.; Gateff, E	R:Roobroek, A.J.; Creemers, J.W.; Pauli, I.G.; Kurzik-Dumke, U.; Rentrop, M.; Gateff, E	A:Title:	Cloning and functional expression of Dfurn2, a subtilisin-like proprotein prc		
A:Reference number:	A3434; MUID:92381036; PMID:1512259	A:Reference number:	A3434; MUID:92381036; PMID:1512259		

A;Accession: A43434
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-1680 <ROE>
 A;Cross-references: GB:MR4375; NID:9157461; PID:9157462
 A;Note: sequence extracted from NCBI backbone (NCBIN:111933, NCBTP:111934)
 C;Keywords: hydrolase; serine proteinase; transmembrane protein
 C;Genetics:
 A;Gene: FlyBase: Fur2
 A;Cross-references: FlyBase: FBgn004598
 C;Superfamily: subtilisin homology
 C;Species: subtilisin homology
 F;409-652;Domain: subtilisin homology <SBT>
 F;418,457,638;Active site: Asp, His, Ser #status predicted
 Query Match Similarity 11.6%; Score 176.5; DB 2; Length 1680;
 Best Local Similarity 28.0%; Pred. No. 0.0003; NID:9157461; PID:9157462
 Matches 60; Conservative 24; Mismatches 77; Indels 53; Gaps 12;
 Qy 3 HLRILISWLIILNFMEXIGSONASGRQRQRMRHMENVSQCGCAGTSDY-NGCLSKPR 61
 Db 1038 HLVWID-LAVCLCPDQYFENs---RNIKTCVP---CEPNCAACDHPREYCTSDH 1087
 Qy 62 LFFPALERIGMKQIGVCLSSCGSGYGRYTRPDINKTKKADDPCTC- -NQNFTCKKKG 119
 Db 1088 LVMHHRK-----CYSACPLDTBT--EDNIAFCHSTCATCNGCPTPQDCITCRSSR 1136
 Qy 120 YLHGGKCLDNCPEGLEANNHMECVSIVCNEVSNAPSPCKKGKRGTRV-R 178
 Db 1137 YAWONKCLISCPDPGFYADKKGECM-----PCOEGCKTC---TSNGVCS 1177
 Qy 179 EIIQHPSAKGNLCPPNTRKCTVORK-CQKG 211
 Db 1178 ECLOMWT-----LAKRDKCIVGSEGSESE 1203

RESULT 6

A39490
 subtilisin-like proprotein convertase (EC 3.4.21.-) PACE4 precursor, splice form A - human
 N;Alternate names: kexin homolog
 C;Species: Homo sapiens (man)
 C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 31-Mar-2000
 C;Accession: A39490
 R;Krieter, M.C.; Tucker, J.B.; Joh, R.; Landsberg, K.E.; Saltman, D.; Barr, P.J.
 DNA Cell Biol. 10, 757-769, 1991
 A;Title: Identification of a second human subtilisin-like protease gene in the fes/fps R
 A;Reference number: A39490; MUID:92075167; PMID:1719556
 A;Accession: A39490
 A;Molecule type: mRNA
 A;Residues: 1-969 <KIE>
 A;Cross-references: GB:MB0482; NID:9189531; PIDN:AAA59998.1; PID:9189532
 C;Genetics:
 A;Gene: GDB:PACE4
 A;Cross-references: GDB:131390; OMIM:167405
 A;Map position: 15q26.1-15q26
 C;Superfamily: subtilisin-like proteinase PACE4; subtilisin homology
 C;Keywords: alternative splicing; hydrolase; serine proteinase
 F;150-969;Protein: serine proteinase PACE4 #status predicted <SIG>
 F;196,434;Domain: subtilisin homology <SBT>
 F;205,246,420;Active site: Asp, His, Ser #status predicted

Query Match Similarity 11.4%; Score 173.5; DB 1; Length 969;
 Best Local Similarity 28.8%; Pred. No. 0.00032; NID:9189531; PID:9189532
 Matches 44; Conservative 16; Mismatches 54; Indels 39; Gaps 7;

Qy 19 YIGSQNAASRGRQRQRMRHMVSQCGGATCSD- -YNGCLSKRPLFALERIGMKIGV 76
 Db 737 YFGDTAARRCR-----CHKGCTCSSLRAATCSCR-RGYF---RHOENNT 779
 Qy 77 CLSSCPSQYGYGRYRDPINKTKCAGCTCFCNK- -NECTKCKSGYFLHLGKCLDNCPEGLE 135
 Db 780 CVCICPAGFYADE- -SOKNCLKCHSKCKVDPKCTVCKEGFSLAROSCTPDCBEGTY 837

Query Match Similarity 11.0%; Score 167.5; DB 2; Length 932;
 Best Local Similarity 24.1%; Pred. No. 0.00076; NID:9169700; PIDN:BA08777.1; PID:9169701
 Matches 52; Conservative 25; Mismatches 78; Indels 61; Gaps 11;

Qy 19 YIGSQNAASRGRQRQRMRHMVSQCGGATCSDY- -GCLSKRPLFALERIGMKIGV 76

RESULT 7

B42225
 probable proprotein convertase (EC 3.4.21.-) 5 precursor - rat
 N;Alternate names: PC5 Precursor
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 23-Feb-1997
 C;Accession: B8825
 R;Lusson, J.; Vieau, D.; Hamelin, J.; Day, R.; Chretien, M.; Seidah, N.G.
 Proc. Natl. Acad. Sci. U.S.A. 90, 6691-6695, 1993
 A;Title: cDNA structure of the mouse and rat subtilisin/kexin-like PC5: a candidate pre
 A;Reference number: A48225; MUID:93342056; PMID:8341687
 A;Accession: B8825
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1915 <LUS>
 C;Cross-references: GB:BL114933
 C;Superfamily: subtilisin-like proteinase PACE4; subtilisin homology
 C;Keywords: duplication; glycoprotein; hydrolase; integrin binding; serine proteinase
 F;1-34;Domain: signal sequence #status predicted <SIG>
 F;35-915;Product: probable proprotein convertase 5 #status predicted <PRO>
 F;117-915;Product: probable proprotein convertase 5 #status experimental <MAT>
 F;164-402;Domain: subtilisin homology <SBT>
 F;173,214,388;Active site: Asp, His, Ser #status predicted
 Query Match Similarity 11.1%; Score 168; DB 2; Length 915;
 Best Local Similarity 26.3%; Pred. No. 0.0007; NID:9169700; PIDN:BA08777.1; PID:9169701
 Matches 49; Conservative 25; Mismatches 54; Indels 58; Gaps 11;

Qy 36 PNVSQ-GCG-----GATCSDYNGLSCPKPLFALARIGMKIGVLLSCPSQYGYR 90
 Db 640 PECBVBGCGDGPBDCTDCLHYKUNTR-----ICVSCEPPCHF--H 682

Qy 91 PDINKCTKCKADCDTFTKPNK- -CTKCKSGFYIH- -LGKCLDNCPEGLEANNHMECVSI 146
 Db 683 ADKCKRCKKCAPNCESEFGSHADQCSCKGYFNEETSSCVACQCEGSSYQDVKNIC- - 739

Qy 147 VHCEVSENPWSPCTKKGKTC- -GFKRGTETRVRREIQQHPSAKGNLCPPNTRKCTVORK 205
 Db 740 -----GKCSNCKTCTGFFHNTB-----CKGGL- -SLQGSRCSV- - 771

Qy 206 KCQKG 211

Db 772 TCEDQ 777

RESULT 8

152527
 PACE4A - mouse (fragment)
 C;Species: Mus musculus (house mouse)
 C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 16-Jul-1999
 C;Accession: 152527
 R;Hosaka, M.; Murakami, K.; Nakayama, K.
 Biomed. Res. 15, 383-390, 1994
 A;Title: PACE4A is a ubiquitous endoprotease that has similar but not identical substrata
 A;Reference number: 152527
 A;Accession: 152527
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-332 <RES>
 A;Cross-references: GB:D50060; NID:9169700; PIDN:BA08777.1; PID:9169701
 C;Superfamily: subtilisin-like proteinase PACE4; subtilisin homology
 F;177-410;Domain: subtilisin homology <SBT>
 Query Match Similarity 11.0%; Score 167.5; DB 2; Length 932;
 Best Local Similarity 24.1%; Pred. No. 0.00076; NID:9169700; PIDN:BA08777.1; PID:9169701
 Matches 52; Conservative 25; Mismatches 78; Indels 61; Gaps 11;

Qy 19 YIGSQNAASRGRQRQRMRHMVSQCGGATCSDY- -GCLSKRPLFALERIGMKIGV 76

Db	700 YFGDAARRCR-----CHKGCCCTGSSPAQCLSCR-RGFY----HHQETNT	742	C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 17-Mar-2000
Qy	77 CLSCLPSGYGTRYPDINKCTCKAKADCDTCNK-NFCTCKSGQFYHLGKCLDNCPEGLE	135	C:Accession: G2428
Db	743 CVALCPAGLYADESQL-CLRCHPSCQKVDEFKCTVCKSGFSLARGSCIPCPBPGY	800	R. Reudelhuber, T.L. submitted to the EMBL Data Library, February 1996
Qy	136 ANNHTMEC-----VS1VHCEVS-----ENNPWPCTKKGKTCGFKRGTERR	176	A:Reference number: H01242
Db	801 FDSELVKGEGCHHTCRTVCGPSRECICAKSFFHQDWKCVPACGE-----GF-----	848	A:Accession: G2428
Qy	177 VREIIIOHPSAKGNICPPNENETRKCTVORKKCGER 212		A:Status: preliminary; translated from GB/EMBL/DDBJ
Db	849 ->---YPEEMPGI--PHKVCRCRCEENCLSCEGSSR 876		A:Molecule type: mRNA
			A:Residues: 1-999 <REU>
			A:Cross-references: EMBL:U49114; NID:91218057; PID:AAA91807.1; PID:91218058
			C:Generics: A:Gene: PC5
			C:Superfamily: subtilisin-like proteinase PACE4; subtilisin homology F;148-386/Domain: subtilisin homology <SBT>
			C:Keywords: hydrolase; serine proteinase F;148-386/Domain: subtilisin homology <SBT>
RESULT 9			
A:8225	subtilisin-like proprotein convertase (EC 3.4.21.-) PC5 precursor - mouse		A:Query Match 10.7%; Score 162; DB 2; Length 899; Best Local Similarity 22.3%; Pred. No. 0.0017; Matches 55; Conservative 33; Mismatches 81; Indels 78; Gaps 12;
N:Alternate names: kexin homolog; serine proteinase PC6			C:Species: Mus musculus (house mouse)
C:Species: Mus musculus (house mouse)			C:Accession: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
R:Accession: J08225; JX0248			R:Accession: J08225; JX0248
Proc. Natl. Acad. Sci. U.S.A. 90, 6891-6895, 1993			A:Reference number: 6891-6895
A:Title: cDNA structure of the mouse and rat subtilisin/kexin-like PC5: a candidate protein			A:Reference number: A48225; MUID:93142056; PMID:8341687
A:Accession: A48225			A:Accession: A48225
A:Status: preliminary			A:Status: preliminary
A:Residues: 1-915 <LUS>			A:Residues: 1-915 <LUS>
R:Nakagawa, T.; Hosaka, M.; Torii, S.; Watanabe, T.; Murokami, K.; Nakayama, K.			R:Nakagawa, T.; Hosaka, M.; Torii, S.; Watanabe, T.; Murokami, K.; Nakayama, K.
J. Biochem. 113, 122-125, 1993			J:Title: Identification and functional expression of a new member of the mammalian kex2-
A:Reference number: JX0248; MUID:93224489; PMID:8463118			A:Accession: JX0248
A:Accession: JX0248			A:Accession: JX0248
A:Residues: 1-915 <NAK>			A:Residues: 1-915 <NAK>
A:Cross-references: DDAS:DJ2619; NID:9220555; PIDN:BA02143.1; PID:9220556			A:Cross-references: DDAS:DJ2619; NID:9220555; PIDN:BA02143.1; PID:9220556
A:Note: the authors translated the codon GGC for residue 915 as Ala			A:Note: the authors translated the codon GGC for residue 915 as Ala
C:Superfamily: subtilisin-like proteinase PACE4; subtilisin homology			C:Superfamily: subtilisin-like proteinase PACE4; subtilisin homology
C:Keywords: duplication; glycoprotein; hydrolase; integrin binding; serine proteinase			C:Keywords: duplication; glycoprotein; hydrolase; integrin binding; serine proteinase
F:1-3/4/Domain: signal sequence #status predicted <PRO>			F:1-3/4/Domain: signal sequence #status predicted <PRO>
F:135-16/Domain: propeptide #status predicted <PRO>			F:135-16/Domain: propeptide #status predicted <PRO>
F:117-915/Product: proprotein convertase PCS #status experimental <MAT>			F:117-915/Product: proprotein convertase PCS #status experimental <MAT>
Query Match 11.0%; Score 167; DB 1; Length 915; Best Local Similarity 27.9%; Pred. No. 0.0081; Matches 51; Conservative 27; Mismatches 53; Indels 52; Gaps 13; F:117-915/Active site: Asp, His, Ser #status predicted			Query Match 11.0%; Score 167; DB 1; Length 915; Best Local Similarity 27.9%; Pred. No. 0.0081; Matches 51; Conservative 27; Mismatches 53; Indels 52; Gaps 13; F:117-915/Active site: Asp, His, Ser #status predicted
Qy	36 PNVPSO-GCGC-GCTCSDYNGLCSKPKRFLFALLERIGMKGQIGVCLASCPGYYGTRYRPI	93	C:Accession: J0148 #sequence_revision 09-May-1997 #text_change 22-Jun-1999
Db	640 PECSEVGCDCGPPGPHCSD--GL---HYYTKL---NNTRICVSSCPGPHY--HADK	685	R:Accession: J0148
Qy	94 NKCTKCKAKACDTERNK- NFCTKCKSGFYIH--LGKCLDNCEBGLANNHMECVSIVHC	149	R:Miranda, L.; Wolf, J.; Pichuantes, S.; Duke, R.; Franzusoff, A.
Db	686 KRCRKKCAPNCEESFGSHQDQCKSGFYIH--LGKCLDNCEBGLANNHMECVSIVHC	171	A:Reference number: J0148; MUID:96353880; PMID:875538
Qy	150 EVSENNPWPCTKKGKTC-GFRGTTTRVREIIQHPSAKGNICPPNENETRKCTVORKKCG	208	A:Contents: CEM T-cell
Db	732 EDIKKAVGCKGKSENCACKIGFNCIE-----CKGGL--SLQGRCSV--TCE	774	A:Accession: J06148
Qy	209 KGE 211		A:Gene: mRNA
Db	775 DQ0 777		A:Residues: 1-915 <MIR>
RESULT 10			A:Cross-references: GB:U56387; NID:91498312; PIDN: AAC50643.1; PID:91498313
G2428	subtilisin-like proprotein convertase (EC 3.4.21.-) 5 precursor - human (fragment)		C:Comment: This protein functions as a soluble enzyme within the secretory pathway. It
C:Species: Homo sapiens (man)			C:Genetics:
C:Gene: PC6A			C:Superfamily: subtilisin-like proteinase PACE4; subtilisin homology
C:Keywords: glycoprotein; hydrolase; serine proteinase			C:Keywords: glycoprotein; hydrolase; serine proteinase
Qy	21 GSQNSRSGRQREMRHPRVSDGCGCATSDYNG--LUSCKERLFLFALLERIGMKGQIGVCL	78	Query Match 10.7%; Score 162; DB 2; Length 915; Best Local Similarity 22.3%; Pred. No. 0.0017; Matches 55; Conservative 33; Mismatches 81; Indels 78; Gaps 12;
Db	679 GHYHADK-KCRKCAPN-----CESCFSHGDQDLSCKGYFL----NFTNSCV	723	C:Accession: J0148 #sequence_revision 09-May-1997 #text_change 22-Jun-1999
Qy	79 SSCPSG-VYGTRYPDINKCTCKAKADCDTCNK-NFCTKCKSGFYHLHGKCLDNCPEGLE	137	C:Species: Homo sapiens (man)

Db 724 TRCPDPSYQDTK---NLCKRKCENCKTCIEFHNCITTECRDGLSLQGSRCVSCEDCRYFN 780
 Qy 138 NHTME-----cys-----iVHCESEW-----NPWPECTK 163
 Db 781 GODCOPCHRFCATCAGAGADGCINCTEGYFMEDGKCVOSCSISVYFDHSSENGYKCKC 840
 Qy 164 GKTG-----GFKGTE-----TRREIOPHSAKGNCPTNPERKCTV 202
 Db 841 DISCLTNCNGPGFKCTSCPSGYLLDGMCOMGAICKDATEESWAGGFCLMLVKQNL--- 898
 Qy 203 ORKKCOK 209
 Db 899 ORKVLOQ 905

RESULT 12

T24322 hypothetical protein R17.3 - *Caenorhabditis elegans*
 C;Species: *Caenorhabditis elegans*
 C;Accession: T24322
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
 R;Barlow, K
 submitted to the EMBL Data Library, March 1997
 A;Accession: T24322
 A;Status: preliminary; Translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-440 <WIL> 292809; PIDN: CAB07269.1; GSPDB: GN00021; CESP: R17.3
 A;Cross-references: EMBL: 292809
 A;Experiment& source: Clone R17
 C;Genetics:

A;Gene: CESP:R17.3

A;Map position: 3
 A;Strains: 36/3; 89/3; 179/2; 241/1; 306/1; 346/2
 A;Supertaxonomy: *Caenorhabditis elegans* hypothetical protein R17.3
 Query Match 10.6%; Score 161; DB 2; Length 440;
 Best Local Similarity 23.7%; Pred. No. 0.002; Mismatches 53; Conservative 30; Gaps 8; Indels 54; Gaps 8;

Db 59 KPRLLFALERIGMKGQIVGLLSSCPSGYIGTRPDINKTKCKADCDCFNFKNFCTKCKSG 118
 Db 172 KPRRHLIIRYSLSKPMLPKVTPSPYLVEENVOPNANLYLESSISBCYCBHC----- 225

Qy 119 FYLHLGKCLDNCPGPLEANHMECVSIVHCVGWNPPSPCTKKGKTCGGKRGTERTRV 178
 Db 226 - -VTLGDC-----CSDYTFVCP--RDCVLTWDSDNTOCADCNGTCG-----GKQKLR 271

Qy 179 EIIQHPSAKANLCPPTNETRKCTVORKKCOKGERGKGRKRKKPNKGESKEAIFDSKS 238
 Db 272 HVIQHARGAACPLKEMTRCVE-----PIKSAADD--- 306

Qy 239 LESKEIPEORENQOQKRK-----VODKOKSVSVSTH 273
 Db 307 ITTVALILDYRHNKTRSKIRRNNTMQLPNVAEKOKKAYCVH 350

RESULT 13

A;9180 Wnt inhibitory factor-1 - human
 C;Species: *Homo sapiens* (man)
 C;Accession: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 18-Feb-2000
 A;Title: A new secreted protein that binds to Wnt proteins and inhibits their activities
 A;Reference number: A59180; MUI: 921555; PMID: 10201374
 A;Accession: A59180
 A;Status: preliminary; not compared with conceptual translation
 A;Molecule type: mRNA
 A;Residues: 1-379 <HSL> A59180; MUI: 921555; PMID: 10201374
 A;Cross-references: GB:AF122922; NID: 94585369; PIDN: AAD25402.1; PID: 94585370

Query Match 10.5%; Score 158.5; DB 2; Length 379;
 Best Local Similarity 24.9%; Pred. No. 0.015;
 Matches 60; Conservative 26; Mismatches 72; Indels 83; Gaps 16;

Db 42 CGCGCAT---CSDVNGC-----LSCKRPLFALERIGMKGQIVCLSS---C 81
 Qy 182 CGCGCNGGFCNERRICECPDFHGRHCEALCPFRMCG-----GLCVTGFCIC 222
 Db 82 PSGYVYTRYPDINKTKC-KADCD-TCFNFKNFCTKCKSGFYLUHGKCLDNCPGEGLEANNH 139
 Qy 233 PPGFY-----VNCDKANCTTCFNGTC---FY-PGKCI-OPPGLEGE--- 271

Db 140 TMECVSIVHCESEWNPSPCTKKGKTCGGKRGTERTRVREIOPHSKGNLCP----- 193
 Db 272 -----QCEISKCP--QPCRNGGKGIG---KSCKCSCKGYQGDLSKPKVCEPGGAHG 318
 Qy 194 -TNETRKCTVORKKCOKGERGKGRKRK-----KPKKGKSEKAIPDSKSLSSKEIP 246
 Db 319 TCEPNKC-----QCEGSHGHGHCNRYEASLHARLPAQAOHQHPSLKAERRDPP 373

Qy 247 E 247
 Db 374 E 374

RESULT 14

Db 15382 gene PACE4 protein - rat
 C;Species: *Rattus norvegicus* (Norway rat)
 C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 22-Jun-1999
 C;Accession: I53282
 R;Johnson, R. C.; Darlington, D. N.; Hand, T. A.; Bloomquist, B. T.; Mains, R. E.
 Endocrinology 115, 1178-1185, 1994
 A;Title: PACE4: a subtilisin-like endoprotease prevalent in the anterior pituitary and
 R;Johnson, R. C.; Darlington, D. N.; Hand, T. A.; Bloomquist, B. T.; Mains, R. E.
 A;Reference number: I53282; MUID: 94349873; PMID: 8070361
 A;Accession: I53282
 A;Status: preliminary; Translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-937 <RES>
 A;Cross-references: GB:L131694; NID: 9496221; PIDN: AAA61987.1; PID: 9496222
 C;Genetics:

A;Gene: PACE4

C;Supertaxonomy: *subtilisin-like proteinase PACE4; subtilisin homology F177-415/Domain: subtilisin*

Query Match 10.3%; Score 156.5; DB 2; Length 937;
 Best Local Similarity 23.6%; Pred. No. 0.004; Mismatches 51; Conservative 24; Gaps 80; Indels 61; Gaps 11;

Db 705 YFGDTAARRR-----CHGCCRCTGKSPTOISCR-RGFR-----JHQETW 747
 Qy 19 YIGSQNSRQRQRMHPNSQGCGCARTCSDN---GCLSKPRLFALERIGMKGQIV 76

Db 777 CLSSCPGSGYTRYPDINKTKC-KADCDTFFNK-NFTCKCKSGFYLUHGKCLDNCPGEGLE 135

Db 748 CTVLCPAGLYADESQRRL---CLRCHPSCKOCVKDFEPEKSTVCKEGFSLARGSCIPDCEPGY 805

Qy 136 ANNHTME-----VSIHCEVS-----ENPNPSPCTKKGKTCGGTERTR 176
 Db 806 FDSELRICGCHCHTCRVCVGPSPREECIHCCKSFHQDWKCVPACGE-----GF----- 853

Qy 177 VREIOPHSAKGNLCPPTNETRKCTVORKKCOKER 212
 Db 854 -----YPEEMPGL-----PKVCKRCDCNCSCEGSSR 881

RESULT 15

A47723 F-spondin precursor - African clawed frog
 C;Species: *Xenopus laevis*
 C;Date: 27-Jun-1994 #sequence_revision 27-Jun-1994 #text_change 17-Nov-2000
 C;Accession: A47723

R: Ruiz, I; Altaba, A.; Cox, C.; Jessell, T.M.; Klar, A.
 Proc. Natl. Acad. Sci. U.S.A. 90, 8368-8272, 1993
 A; Title: Ectopic neural expression of a floor plate marker in frog embryos injected with
 A; Reference number: A47723; MUID:93376785; PMID:8367492
 A; Accession: A47723
 A; Status: preliminary
 A; Molecule type: mRNA
 A; Residues: 1-803 <RUI>
 C; Cross-references: GB:09123; NID:9409244; PID:AAA19105_1; PID:9409245
 C; Superfamily: F-spondin; thrombospondin type 1 repeat homology <THR2>
 F; 35-489/Domain: thrombospondin type 1 repeat homology <THR1>
 F; 607-662/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 10.3%; Score 156; DB 2; Length 803;
 Best Local Similarity 24.6%; Pred. No. 0039; Mismatches 91; Indels 52; Gaps 12;
 Matches 58; Conservative 35; Mismatches 91; Indels 52; Gaps 12;
 Qy 38 VSQGCQGGCATCSDYNGCLSKPRLFALLERIGMMQIGVCLSSCPs-GYyGTRYPPDINK 95
 Db 544 VNEECPSCLIVTWAENBECs----ATCGMGKRRHMKOPTPGSMKADTEVEK 598
 Qy 96 C--TKK-----ADCDTCFNKNFCTKCSGFGYLHLGKCLDNCPPGLEANN-- 138
 Db 599 CMMPECHTIPCVLSPWSEWSDCSvTCKGTRTRQK--MLKSSSELDNCNEELEKQVEK 655
 Qy 139 -HTMECVSIVHCEVSEWNPCTKKTCGFKRGMETRVEI1QHPSAKGNLCPPNET 197
 Db 656 CMLPCEP--1SCETEWNSWSEC--NKSQG--KSHMIRTRMTMEQFGGavCETVQR 708
 Qy 198 RKCTVORKKKCQKGRKGKKRERKRRKKRNKGESKEAIPDSKLESSKETPEORENKO 253
 Db 709 KKCRK--RKQK---SSONERRHLK-----DAREKRSERKSEKIKEDSGQ 747

Search completed: May 6, 2003, 14:53:03
 Job time : 20 secs

PR 28-JUN-2000; 2000US-215733P.
 PR 09-JAN-2001; 2001US-0757562.
 PR 05-FEB-2001; 2001US-266614P.
 XX
 XX
 (HYSE-) HYSEQ INC.
 PA (KIRI) KIRIN BEER KK.
 XX
 PI Tang TY, Labat I, Tillinghast JS, Sinku A, Liu C, Drmanac RT;
 PI Stache-Crain B, Dickson M, Mize NK, Nishikawa M;
 XX
 DR WPI; 2001-657166/75.
 DR N-PSDB; AAD21725.
 XX
 Novel stem cell growth factor like polypeptides and polynucleotides for identifying modulators useful for treating diseases such as Alzheimer's disease, cancer, rheumatoid arthritis, osteoporosis
 PT Claim 28; Page 214-215; 232pp; English.
 CC The patent discloses novel stem cell growth factor-like proteins and polynucleotides encoding them. Proteins of the invention are also known as supporting factor for the proliferation of stem cells (SCR-1). Stem cell growth factor-like proteins are useful for supporting proliferation or survival of a stem cell or germ cell which is preferably primordial germ cell, germ line stem cell, embryonic stem cell, haematopoietic stem cell, haematopoietic progenitor cell, pluripotent cell or totipotent cell. The haematopoietic progenitor cell cultured using stem cell growth factor-like proteins can replace as a graft for the bone marrow transplantation or cord blood transplantation for treating a variety of diseases such as immunodeficiency syndrome, chronic granulomatous disease, duplicated immunodeficiency syndrome, agammaglobulinaemia, Wiskott-Aldrich syndrome, acquired immune deficiency syndrome (AIDS), thalassaemia, haemolytic anaemia due to enzyme defect, congenital anaemia such as sickle cell anaemia, Gaucher's disease, lysosomal storage diseases such as mucopolysaccharidosis, adrenal white matter degeneration, a variety of cancer and tumours. Proteins of the invention are useful for treating diseases such as Parkinson's disease, Alzheimer's disease and other neurodegenerative diseases, thrombocytopaenia, immune deficiencies and disorders such as severe combined immunodeficiency (SCID) and autoimmune disorders such as severe multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, and autoimmune pulmonary inflammation. Sequences of the invention are also useful in gene therapy. The present sequences is stem cell growth factor-like protein from human.
 CC Sequence 273 AA;
 SQ Query March 100.0%; Score 1515; DB 22; Length 273;
 Best Local Similarity 100.0%; Pred. No. 3e-112; Mismatches 273; Conservative 0; Indels 0; Gaps 0;
 Matches 1
 OY 1 MCHHLRISMFLIINFMEYIGSOMASGRRRPARRMMNVSOGCOGATCSYNGCISK 60
 Db 1 MGHLRLISFLIINFMEYIGSOMASGRRRPARRMMNVSOGCOGATCSYNGCISK 60
 OY 61 RLFALERIGKQGCVLSSCPGSKYGRYDINCKTCKAPEDTCRKNICKTCKGKFY 120
 Db 61 RLFALERIGKQGCVLSSCPGSKYGRYDINCKTCKAPEDTCRKNICKTCKGKFY 120
 XX
 OY 121 LHGKCLDNCPGEGCLEANHTMECVSEWNPWSCTKCKTCGKAGTERPREI 180
 Db 121 LHGKCLDNCPGEGCLEANHTMECVSEWNPWSCTKCKTCGKAGTERPREI 180
 OY 181 IQHPSAKGNCPPNEMTRCTYQKKQGKGKGKGRKRNKPKNSKAIPIPKSL 240
 Db 181 IQHPSAKGNCPPNEMTRCTYQKKQGKGKGKGRKRNKPKNSKAIPIPKSL 240
 OY 241 SSKEIPEORENKOQKQRKVQDKQSKSYVTH 273
 Db 241 SSKEIPEORENKOQKQRKVQDKQSKSYVTH 273
 CC
 AAEE13150
 ID AAEE13150 standard; Protein; 272 AA.
 XX
 AAEE13150;
 XX
 28-JAN-2002 (first entry)
 XX
 DE Human stem cell growth factor-like protein #2.
 XX Human; stem cell growth factor-like protein; antiinflammatory; nootropic; neuroprotective; valnary; cytostatic; anticonvulsant; immunostimulant; vasotrophic; virucide; dermatological; tranquilliser; cerebroprotective; osteopathic; immunodeficiency syndrome; chronic granulomatous disease; duplicated immunodeficiency syndrome; Wiskott-Aldrich syndrome; AIDS; acquired immune deficiency syndrome; agammaglobulinaemia; thalassaemia; Gaucher's disease; lysosomal storage disease; mucopolysaccharidosis; adrenal white matter degeneration; anaemia; neurodegenerative disease; Parkinson's disease; Alzheimer's disease; thrombocytopaenia; SCID; severe combined immunodeficiency; immune disorder; autoimmune disorder; multiple sclerosis; systemic lupus erythematosus; rheumatoid arthritis; autoimmune pulmonary inflammation; cancer; tumour; gene therapy; SCR-1; supporting factor for the proliferation of stem cell.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..21
 FT FT /label= signal_peptide
 FT FT 22..272
 FT FT /note= "Human nature stem cell growth factor-like protein"
 XX
 PN W0200177169-A2.
 XX
 PD 18-OCT-2001.
 XX
 PP 05-APR-2001; 2001WO-US11208.
 XX
 PR 05-APR-2000; 2000US-0543774.
 PR 28-JUN-2000; 2000US-215733P.
 PR 09-JAN-2001; 2001US-0757562.
 PR 05-FEB-2001; 2001US-266614P.
 XX
 PA (HYSE-) HYSEQ INC.
 PA (KIRI) KIRIN BEER KK.
 XX
 PI Tang TY, Labat I, Tillinghast JS, Sinku A, Liu C, Drmanac RT;
 PI Stache-Crain B, Dickson M, Mize NK, Nishikawa M;
 XX
 DR WPI; 2001-657166/75.
 DR N-PSDB; AAD21724.
 XX
 PT Novel stem cell growth factor like polypeptides and polynucleotides for identifying modulators useful for treating diseases such as Alzheimer's disease, cancer, rheumatoid arthritis, osteoporosis
 PT Claim 28; Page 211-212; 232pp; English.
 XX
 CC The patent discloses novel stem cell growth factor-like proteins and polynucleotides encoding them. Proteins of the invention are also known as supporting factor for the proliferation of stem cells (SCR-1). Stem cell growth factor-like proteins are useful for supporting proliferation or survival of a stem cell or germ cell which is preferably primordial germ cell, germ line stem cell, embryonic stem cell, haematopoietic stem cell, haematopoietic progenitor cell, pluripotent cell or totipotent cell. The haematopoietic progenitor cell cultured using stem cell growth factor-like proteins can replace as a graft for the bone marrow transplantation or cord blood transplantation for treating a variety of diseases such as immunodeficiency syndrome, chronic granulomatous disease, duplicated immunodeficiency syndrome, agammaglobulinaemia, Wiskott-Aldrich syndrome, acquired immune deficiency syndrome (AIDS), thalassaemia, haemolytic anaemia due to enzyme defect, congenital anaemia such as sickle cell anaemia, Gaucher's disease, lysosomal

storage diseases such as mucopolysaccharidosis, adrenal white matter degeneration, a variety of cancer and tumours. Proteins of the invention are useful for treating diseases such as Parkinson's disease, Alzheimer's disease and other neurodegenerative diseases, thrombocytopathy, immune deficiencies and disorders such as severe combined immunodeficiency (SCID) and autoimmune disorders such as multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, and autoimmune pulmonary inflammation. Sequences of the invention are also useful in gene therapy. The present sequence is stem cell growth factor-like protein from human.

CC Sequence 272 AA;

Query Match 99.3%; Score 1505; DB 22; Length 272; Best Local Similarity 100.0%; Pred. No. 2.3e-111; Mismatches 0; Conservative 0; Indels 0; Gaps 0;

QY 3 HLRILISWLFILNPMEMYIGSONASRRQRMRMPNVSOCGCGCATSDYNGLCKSPKRL 62
Db 2 HLRILISWLFILNPMEMYIGSONASRRQRMRMPNVSOCGCGCATSDYNGLCKSPKRL 61
QY 63 FFALERIGMKQIGVCLSCPGSYGTRYPDINKCTKCKRADCDTCFCNFKCNSGCFKSPKRL 62
Db 62 FFALERIGMKQIGVCLSCPGSYGTRYPDINKCTKCKRADCDTCFCNFKCNSGCFKSPKRL 61
QY 123 LGKCLDNCPGEGLEAHNMTCVSIWCEVEWNPNSPCTKKGKTCFKGTRTTRRILQ 182
Db 122 LGKCLDNCPGEGLEAHNMTCVSIWCEVEWNPNSPCTKKGKTCFKGTRTTRRILQ 182
QY 183 HPSAKGNLCPPTNETKCTVORKCOKGERGKKRERKRKPNKGSKESKAIPDSKLSS 242
Db 182 HPSAKGNLCPPTNETKCTVORKCOKGERGKKRERKRKPNKGSKESKAIPDSKLSS 241
QY 243 KEIPQRENQKQOKKKVQDKQSKVSVTH 273
Db 242 KEIPQRENQKQOKKKVQDKQSKVSVTH 272

RESULT 3

ID AAE13168

AC AAE13168;

XX DT 28-JAN-2002 (first entry)

DE Human stem cell growth factor-like protein #4.

Human; stem cell growth factor-like protein; antiinflammatory; nontropic; neuroprotective; vulnerar; cytostatic; anticonvulsant; immunostimulant; osteoprotic; virucide; dermatologic; tranduillise; cerebroprotective; duplicated immunodeficiency syndrome; Wiskott-Aldrich syndrome; AIDS; acquired immune deficiency syndrome; agammaglobulinaemia; thalassaemia; Gaucher's disease; lysosomal storage disease; mucopolysaccharidosis; adrenal white matter degeneration; anaemia; neurodegenerative disease; Parkinson's disease; Alzheimer's disease; thrombocytopaenia; SCID; severe combined immunodeficiency; immune disorder; autoimmune disorder; multiple sclerosis; systemic lupus erythematosus; rheumatoid arthritis; autoimmune pulmonary inflammation; gene therapy; SCR-1; supporting factor for the proliferation of stem cell. Homo sapiens.

FH KEY location/qualifiers

1..21 /label= Signal_peptide
22..272 /note= "Human mature stem cell growth factor-like protein"

PN WO200177169-A2.

XX

PD 18-OCT-2001.
XX PF 05-APR-2001; 2001WO-US11208.

XX PR 05-APR-2000; 2000US-0543774.
PR 02-JUN-2000; 2000US-215733P.
PR 09-JAN-2001; 2001US-0757562.
PR 05-FEB-2001; 2001US-266514P.

XX PA (HYSE-) HYSEQ INC.
PA (KIRIN BEER KK.)

XX PI Tang TW, Labat I, Tillinghast JS, Sinku A, Liu C, Drmanac RT;
PI Strache-Crain B, Dickson M, Mize NK, Nishikawa M;

XX DR N-PSDB; ABD21728.

PT Novel stem cell growth factor like polypeptides and polynucleotides for identifying modulators useful for treating diseases such as Alzheimer's disease, cancer, rheumatoid arthritis, osteoporosis

PT PS XX

Claim 28; Page 226-227; 22pp; English.

The patent discloses novel stem cell growth factor-like proteins and polynucleotides encoding them. Proteins of the invention are also known as supporting factor for the proliferation of stem cells (SCR-1). Stem cell growth factor-like proteins are useful for supporting proliferation or survival of a stem cell or germ cell which is preferably primordial germ cell, germ line stem cell, embryonic stem cell, haemopoietic stem cell, haematopoietic progenitor cell, pluripotent cell or totipotent cell. The haematopoietic progenitor cell cultured using stem cell growth factor-like proteins can replace as a graft for the bone marrow transplantation or cord blood transplantation for treating a variety of diseases such as immunodeficiency syndrome, chronic granulomatous disease, duplicated immunodeficiency syndrome, agammaglobulinaemia, Wiskott-Aldrich syndrome, acquired immune deficiency syndrome (AIDS), thalassaemia, haemolytic anaemia due to enzyme defect, congenital anaemia such as sickle cell anaemia, Gaucher's disease, lysosomal storage diseases such as mucopolysaccharidosis, adrenal white matter degeneration, a variety of cancer and tumours. Proteins of the invention are useful for treating diseases such as Parkinson's disease, Alzheimer's disease and other neurodegenerative diseases, thrombocytopaenia, immune deficiencies and disorders such as severe combined immunodeficiency (SCID) and autoimmune disorders such as multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, and autoimmune pulmonary inflammation. Sequences of the invention are also useful in gene therapy. The present sequence is stem cell growth factor-like protein from human.

XX Sequence 272 AA;

Query Match 99.3%; Score 1505; DB 22; Length 272; Best Local Similarity 100.0%; Pred. No. 2.3e-111; Mismatches 0; Conservative 0; Indels 0; Gaps 0;

QY 3 HLRILISWLFILNPMEMYIGSONASRRQRMRMPNVSOCGCGCATSDYNGLCKSPKRL 62

Db 2 HLRILISWLFILNPMEMYIGSONASRRQRMRMPNVSOCGCGCATSDYNGLCKSPKRL 61

QY 63 FFALERIGMKQIGVCLSCPGSYGTRYPDINKCTKCKRADCDTCFCNFKCNSGCFKSPKRL 62

Db 62 FFALERIGMKQIGVCLSCPGSYGTRYPDINKCTKCKRADCDTCFCNFKCNSGCFKSPKRL 61

QY 123 LGKCLDNCPGEGLEAHNMTCVSIWCEVEWNPNSPCTKKGKTCFKGTRTTRRILQ 182

Db 122 LGKCLDNCPGEGLEAHNMTCVSIWCEVEWNPNSPCTKKGKTCFKGTRTTRRILQ 181

QY 183 HPSAKGNLCPPTNETKCTVORKCOKGERGKKRERKRKPNKGSKESKAIPDSKLSS 242

Db 182 HPSAKGNLCPPTNETKCTVORKCOKGERGKKRERKRKPNKGSKESKAIPDSKLSS 241

QY 243 KEIPQRENQKQOKKKVQDKQSKVSVTH 273

XX

Db	242	KEIPEORENQOKRKVQDKOKSVSVSTVH	272	Db	2	HLRLISWLFILNPMVEYIGSONASRGRRQRMRPHNVSOGCGCATCSDYNGCLSKPRL	61
Qy	63	FFALERIGMKGQIGVCLSCPSGCGYGRYDPIKCTKACADCTCFNNFCFTKCKSGFYIH	122	Qy	63	FFALERIGMKGQIGVCLSCPSGCGYGRYDPIKCTKACADCTCFNNFCFTKCKSGFYIH	122
Db	62	FFALERIGMKGQIGVCLSCPSGCGYGRYDPIKCTKACADCTCFNNFCFTKCKSGFYIH	121	Db	62	FFALERIGMKGQIGVCLSCPSGCGYGRYDPIKCTKACADCTCFNNFCFTKCKSGFYIH	121
RESULT 4	XX	XX	XX	Qy	123	LGKCDNCPGLENNTMECVSEWNPSEPTKKGKTCFKRTEYTRREIQ	182
ID	AM78328	standard; Protein; 272 AA.		Db	122	LGKCDNCPGLENNTMECVSEWNPSEPTKKGKTCFKRTEYTRREIQ	181
XX	XX	XX		Qy	183	HPSAGNLLCPTPTETRKCTVORKCOKGKGKGRERKKPNKSKEAPDSKLSS	242
AC	AAM78328;			Db	182	HPSAGNLLCPTPTETRKCTVORKCOKGKGKGRERKKPNKSKEAPDSKLSS	241
DT	06-NOV-2001	(first entry)		Qy	243	KEIPEORENQOKRKVQDKOKSVSVSTVH	273
XX	XX	XX		Db	242	KEIPEORENQOKRKVQDKOKSVSVSTVH	272
DE	human protein SEQ ID NO 990.			Qy	244	KEIPEORENQOKRKVQDKOKSVSVSTVH	274
XX	Human: cytokine; cell proliferation; cell differentiation; gene therapy; vaccine; peptide therapy; stem cell growth factor; haematopoiesis; tissue growth factor; immunomodulatory; cancer; leukaemia; nervous system disorder; arthritis; inflammation.			Db	245	KEIPEORENQOKRKVQDKOKSVSVSTVH	275
OS	Homo sapiens.			Qy	246	KEIPEORENQOKRKVQDKOKSVSVSTVH	276
XX	XX			Db	247	KEIPEORENQOKRKVQDKOKSVSVSTVH	277
PN	W0200157190-A2.			Qy	248	KEIPEORENQOKRKVQDKOKSVSVSTVH	278
XX	XX			Db	249	KEIPEORENQOKRKVQDKOKSVSVSTVH	279
PD	09-AUG-2001.			Qy	250	KEIPEORENQOKRKVQDKOKSVSVSTVH	280
XX	XX			Db	251	KEIPEORENQOKRKVQDKOKSVSVSTVH	281
PF	05-FEB-2001; 2001WO-US04098.			Qy	252	KEIPEORENQOKRKVQDKOKSVSVSTVH	282
XX	XX			Db	253	KEIPEORENQOKRKVQDKOKSVSVSTVH	283
PR	03-FEB-2000; 2000US-0436914.			Qy	254	KEIPEORENQOKRKVQDKOKSVSVSTVH	284
PR	27-APR-2000; 2000US-0450875.			Db	255	KEIPEORENQOKRKVQDKOKSVSVSTVH	285
PR	20-JUN-2000; 2000US-0598075.			Qy	256	KEIPEORENQOKRKVQDKOKSVSVSTVH	286
PR	19-JUL-2000; 2000US-0620325.			Db	257	KEIPEORENQOKRKVQDKOKSVSVSTVH	287
PR	01-SEP-2000; 2000US-0649325.			Qy	258	KEIPEORENQOKRKVQDKOKSVSVSTVH	288
PR	15-SEP-2000; 2000US-0663561.			Db	259	KEIPEORENQOKRKVQDKOKSVSVSTVH	289
PR	30-OCT-2000; 2000US-0693325.			Qy	260	KEIPEORENQOKRKVQDKOKSVSVSTVH	290
PR	30-NOV-2000; 2000US-0728422.			Db	261	KEIPEORENQOKRKVQDKOKSVSVSTVH	291
XX	(HYSE-)	HYSEQ INC.		Qy	262	KEIPEORENQOKRKVQDKOKSVSVSTVH	292
XX	XX			Db	263	KEIPEORENQOKRKVQDKOKSVSVSTVH	293
PT	Tang YI, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Gao Y, Ma Y;			Qy	264	KEIPEORENQOKRKVQDKOKSVSVSTVH	294
PT	PT	Zhao Q, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;		Db	265	KEIPEORENQOKRKVQDKOKSVSVSTVH	295
PT	PT	Xue AJ, Yang Y, Wejhrman T, Goodrich R;		Qy	266	KEIPEORENQOKRKVQDKOKSVSVSTVH	296
XX	XX	WPI; 2001-476293/51.		Db	267	KEIPEORENQOKRKVQDKOKSVSVSTVH	297
DR	N-PSDB; AAK51461.			Qy	268	KEIPEORENQOKRKVQDKOKSVSVSTVH	298
XX	XX			Db	269	KEIPEORENQOKRKVQDKOKSVSVSTVH	299
PT	Nucleic acids encoding polypeptides with cytokine-like activities, useful in diagnosis and gene therapy.			Qy	270	KEIPEORENQOKRKVQDKOKSVSVSTVH	300
XX	XX			Db	271	KEIPEORENQOKRKVQDKOKSVSVSTVH	301
PS	Claim 20; Page 3214-3215; 6221pp; English.			Qy	272	KEIPEORENQOKRKVQDKOKSVSVSTVH	302
XX	XX			Db	273	KEIPEORENQOKRKVQDKOKSVSVSTVH	303
CC	The invention relates to polynucleotides (AAK51456-AAK53435) and the encoded polypeptides (AM78323-AM80321) that exhibit activity elating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haemopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activitin/inhibin activity and may be useful in the diagnosis and treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation.			Qy	274	KEIPEORENQOKRKVQDKOKSVSVSTVH	304
CC	Note: Records for SEQ ID NO 2110 (AAK5281), 2111 (AAK5282) and 3666 (AM80020) are omitted as the relevant pages from the sequence listing were missing at the time of publication.			Db	275	KEIPEORENQOKRKVQDKOKSVSVSTVH	305
XX	XX			Qy	276	KEIPEORENQOKRKVQDKOKSVSVSTVH	306
Sequence	Sequence	272 AA;		Db	277	KEIPEORENQOKRKVQDKOKSVSVSTVH	307
Qy	Query Match	99.3%; Score 1505; DB 22; Length 272;		Qy	278	KEIPEORENQOKRKVQDKOKSVSVSTVH	308
Qy	Best Local Similarity	100.0%; Pred. No. 2.3e-11; Mismatches 0; Indels 0; Gaps 0;		Db	279	KEIPEORENQOKRKVQDKOKSVSVSTVH	309
Qy	Matches	271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		Qy	280	KEIPEORENQOKRKVQDKOKSVSVSTVH	310
Qy	3	HRLISWLFILNPMVEYIGSONASRGRRQRMRPHNVSOGCGCATSDYNGCLSKPRL	62	Qy	281	HRLISWLFILNPMVEYIGSONASRGRRQRMRPHNVSOGCGCATSDYNGCLSKPRL	62

Db	2	HLRLISWLFLILNFMEYIGQSNAQRGRQRMRPHNTSGOCGGCATSDYNGCSCCKPRL
Qy	63	FPFALEIIGMKGQIGVCLSSPESPGYGYTRPDDINCTKQKTCADDTCEFNKFCTKCSGQFYH
Db	62	FFALERIGMKGQIGVCLSSPESPGYGYTRPDDINCTKQKTCADDTCEFNKFCTKCSGQFYH
Qy	123	LGKCLDNCPEGLEANHTHCVSIVHCEVSERNPWSCTKQGTCFKRTEVREIQLH
Db	122	LGKCLDNCPEGLEANHTHCVSIVHCEVSERNPWSCTKQGTCFKRTEVREIQLH
Qy	183	HPSAKGNLCPPTNERKTCVORKKQOKBERGKKGGERKRPKPNKGSEKIPDSKLESS
Db	182	HPSAKGNLCPPTNETRKTCVORKKQOKGERGKKGGERKRPKPNKGSEKAPDSKLESS
Qy	243	KELPEORENQOKQKRVQDKQKSVSVTH 273
Db	242	KELPEORENQOKQKRVQDKQKSVSVTH 272
RESULT 6		
ID	AAB13163	standard; Protein; 265 AA.
XX	AAE13163;	
AC		
XX		
DT	28-JAN-2002	(first entry)
DE	Human secreted protein from clone DA228_5.	
XX	Human; stem cell growth factor-like protein; antiinflammatory; nootropics; neuroprotective; vulnerability; cytostatic; anticonvulsant; immunostimulant; vasotropics; viceide; dermatologic; tranquilizer; cerebroprotective; osteopathic; immunodeficiency syndrome; chronic granulomatous disease; duplicated immunodeficiency syndrome; Wiskott-Aldrich syndrome; AIDS; acquired immune deficiency syndrome; agammaglobulinemia; thalassemia; Gaucher's disease; lysosomal storage disease; mucopolysaccharidoses; adrenale white matter degeneration; anaemia; neurodegenerative disease; parkinson's disease; Alzheimer's disease; thrombocytopoenia; SCID; severe combined immunodeficiency; immune disorder; autoimmune disorder; multiple sclerosis; systemic lupus erythematosus; rheumatoid arthritis; autoimmune pulmonary inflammation; cancer; tumour; gene therapy; SC-1; supporting factor for the proliferation of stem cell; secreted protein.	
XX	OS Homo sapiens.	
XX	PN	
XX	W020017169-A2.	
PD	18-OCT-2001.	
XX		
PF	05-APR-2001; 2001WO-US11208.	
XX		
PR	05-APR-2000; 2000US-0543774.	
PR	28-JUN-2000; 2000US-215733P.	
PR	09-JAN-2001; 2001US-075562.	
PR	05-FEB-2001; 2001US-266614P.	
XX		
PA	(HYSE-) HYSEQ INC.	
PA	(KIRI) KIRIN BEER KK.	
XX		
PT	Tang TY, Labar I, Tillinghast JS, Sinku A, Liu C, Drmanac RT;	
PT	Stacine-Crain B, Dickson M, Mize NK, Nishikawa M;	
PS	Disclosure; Fig 3; 232PP; English.	
XX		
CC	The patent discloses novel stem cell growth factor-like proteins and polynucleotides encoding them. Proteins of the invention are also known	
CC	Novel stem cell growth factor like polypeptides and polynucleotides for identifying modulators useful for treating diseases such as Alzheimer's disease, cancer, rheumatoid arthritis, osteoporosis -	
CC	Disclosure; Fig 3; 232PP; English.	
DR	WPI; 2001-65716/75.	
XX		

as supporting factor for the proliferation of stem cells (SCR-1); Stem cell growth factor-like proteins are useful for supporting proliferation or survival of a stem cell or germ cell which is preferably primordial germ cell, germ line stem cell, embryonic stem cell, hematopoietic stem cell, hematopoietic progenitor cell, pluripotent cell or totipotent cell. The haematopoietic progenitor cell, cultured using stem cell growth factor-like proteins can replace as a graft for the bone marrow transplantation or cord blood transplantation for treating a variety of diseases such as immunodeficiency syndrome, chronic granulomatous disease, duplicated immunodeficiency syndrome, agammaglobulinaemia, Wiskott-Aldrich syndrome, acquired immune deficiency syndrome (AIDS), thalassaemia, haemolytic anaemia due to enzyme defect, congenital anaemia such as sickle cell anaemia, Gaucher's disease, lysosomal storage diseases such as mucopolysaccharidosis, adrenal white matter degeneration, a variety of cancer and tumours. Proteins of the invention are useful for treating diseases such as Parkinson's disease, Alzheimer's disease and other neurodegenerative diseases, thrombocytopaenia, immune deficiencies and disorders such as severe combined immunodeficiency (SCID) and autoimmune disorders such as multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, and autoimmune pulmonary inflammation. Sequences of the invention are also useful in gene therapy. The present sequence is human secreted protein from clone DA228-6.

XX

Sequence	265 AA;
Query Match	97.1%; Score 1472; DB 22; Length 265;
Best Local Similarity	100.0%; Pred. No. 9e-109;
Matches	264; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	3 HRLILWLFITLNFMEYIGSSONASRGRQRKRPMPNSQGCGCATCSDYNGCISCKPRL 62
Db	2 HRLISLWLFITLNFMEYIGSSONASRGRQRKRPMPNSQGCGCATCSDYNGCISCKPRL 61
Qy	63 FFALEIGMIGKQIGVGLSSERSSRGYGRTPDINKCTCKADDDTCENKNFCTCKSGPYIH 122
Db	62 FFALERIGMIGQIGVGLSSERSSPGYIGTRTPDINKCTCKADDDTCENKNFCTCKSGPYIH 121
Qy	123 LGKCLNQCPGLEYAQNHTWCVSIVHCEVSINPWPSCPCKGKGKTCFGKSGTETRYEIQ 182
Db	122 LGKCLDNCPGLEYAQNHTWCVSIVHCEVSINPWPSCPCKGKGKTCFGKRGETRYEIQ 181
Qy	183 HPSAKGNLCPPTNETRKCTVORKKCKGERKGERKGERKPKRNKGESKEIPDSKLESS 242
Db	182 HPSAKGNLCPPTNETRKCTVORKKCKGERKGGKGERERKPKRNKGESKEAPDSKLESS 241
Qy	243 KEIPEORENQOQKGRKVQDKQKS 266
Db	242 KEIPEORENQOQKGRKVQDKQKS 265

RESULT 7

ID	AAW85607 standard; Protein: 292 AA.
XX	AAW85607;
AC	AAW85607;
XX	02-MAR-1999 (first entry)
XX	Secreted protein clone da228_6.
XX	Clone; secreted protein; protein factor; cytokine; lymphokine; KW interferon; colony stimulating factor; Csf; interleukin; cloning; KW tumour invasion; tumour suppression; immune boosting.
XX	OS Homo sapiens.
XX	W09849302-A1.
XX	PD 05-NOV-1998.
XX	PF 24-APR-1998; 98WO-US08336.
XX	PN
XX	PD
XX	PF
XX	XX

PR 23-APR-1998; 98US-0065125.
 PR 25-APR-1997; 97US-0045296.
 XX (GEMY) GENETICS INST INC.
 XX Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
 PI Racie LA, Spaulding V, Treacy M;
 XX DR WPI; 1999-024059/02.
 DR N-PSDB; AAV83133.

PT New polynucleotides encoding secreted human proteins • are derived from human foetal brain, adult brain, adult blood or placenta cDNA libraries, useful, e.g. as potential immunomodulators

PT XX
 PS XX
 PT XX
 XX
 CC The nucleotide sequence (NS) of the full-length protein-coding sequence of clones csi54 (AAV8132), da2286 (AAV83133), du1105 (AAV8134), eh801 (AAV83135), er3691 (AAV83136), flm1235 (AAV83137), fml1105 (AAV8138) or fr4312 (AAV8139), (all clones are deposited as ATCC 98451) and the proteins they encode are predicted to have biological activities which would make them suitable for treating, preventing or ameliorating medical conditions in human and animals for example, tumour suppression/invasion activity, immune system boosting activity. The polynucleotides are also believed to be useful for gene therapy.

XX SQ Sequence 292 AA;
 XX Query March 97.1%; Score 1472; DB 20; Length 292;
 Best Local Similarity 100.0%; Pred. No. 1e-108; Mismatches 0; Indels 0; Gaps 0;
 Matches 264; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 HRLISLWLFILNFMEMYIGSONASRRQRMRHMHNVSQCGCATSDYNGLSCKRL 62
 Db 2 HRLISLWLFILNFMEMYIGSONASRRQRMRHMHNVSQCGCATSDYNGLSCKRL 61

Qy 63 FFLALRIGMKQIGVCLSCPSQGYGRYPDINKCTKCKADCTCFNFKNFTCKCSGFYIH 122
 Db 62 FFLALRIGMKQIGVCLSCPSQGYGRYPDINKCTKCKADCTCFNFKNFTCKCSGFYIH 121

Qy 123 LGKCLDNCPGEGLENNHMECVSIVHCESEWNWNSPCTKKTCEPKGRMETRVEIQ 182
 Db 122 LGKCLDNCPGEGLENNHMECVSIVHCESEWNWNSPCTKKTCEPKGRMETRVEIQ 181

Qy 183 HPSAKNLCPPTNETRKCTVORKKCKGKGRRERKKPKNGESKEAIPDSKLESS 242
 Db 182 HPSAKNLCPPTNETRKCTVORKKCKGKGRRERKKPKNGESKEAIPDSKLESS 241

Qy 243 KEIPREOREKKQOKRKVQDKKS 266
 Db 242 KEIPREOREKKQOKRKVQDKKS 265

RESULT 8

AAE13170 ID AAE13170 standard; Protein: 292 AA.
 ID AAE13170
 XX AC AAE13170;
 XX DT 28-JAN-2002 (first entry)
 DE Human SCR-1 related protein.

XX Human: stem cell growth factor-like protein; antiinflammatory; nootropic; neuroprotective; vulnerary; cytoprotective; anticonvulsant; immunostimulant; vasotocpic; vasoconstrictive; dermatological; tranquiliser; cerebroprotective; osteopathic; immunodeficiency syndrome; chronic granulomatous disease; duplicated immunodeficiency syndrome; Wiskott-Aldrich syndrome; AIDS; acquired immune deficiency syndrome; adammaglobulinaemia; thalassemia; Kw Gaucher's disease; lysosomal storage disease; mucopolysaccharidosis; adrenal white matter degeneration; anaemia; neurodegenerative disease;

KW Parkinson's disease; Alzheimer's disease; thrombocytopaenia; SCID; severe combined immunodeficiency; immune disorder; autoimmune disorder; multiple sclerosis; systemic lupus erythematosus; rheumatoid arthritis; KW autoimmune pulmonary inflammation; cancer; tumour; gene therapy; SCR-1; supporting factor for the proliferation of stem cell. Unidentified.

XX OS DR W0200177169-A2.
 XX PN XX
 PD 18-OCT-2001.
 XX PT XX
 PR 05-APR-2001; 2001WO-US11208.
 PR 05-APR-2000; 2000US-0543774.
 PR 28-JUN-2000; 2000US-215731P.
 PR 09-JAN-2001; 2001US-0757562.
 PR 05-FEB-2001; 2001US-266614P.
 PA (HYSE-) HYSEQ INC.
 PA (KIRI) KIRIN BEER KK.
 PA Tang TY, Labat I, Tillinghast JS, Sinku A, Liu C, Drmanac RT;
 PI Schae-Crain B, Dickson M, Mize NK, Nishikawa M;
 PI DR WPI; 2001-657166/75.
 DR N-PSDB; AAD21740.

XX PT Novel stem cell growth factor like polypeptides and polynucleotides for identifying modulators useful for treating diseases such as Alzheimer's disease, cancer, rheumatoid arthritis, osteoporosis

XX PS XX
 PT Claim 27; Page 231-232; 232pp; English.

CC The patent discloses novel stem cell growth factor-like proteins and polynucleotides encoding them. Proteins of the invention are also known and supporting factor for the proliferation of stem cells (SCR-1). Stem cell growth factor-like proteins are useful for supporting proliferation or survival of a stem cell or germ cell which is preferably primordial germ cell, germ line stem cell, embryonic stem cell, haematopoietic stem cell, haematopoietic progenitor cell, pluripotent cell or totipotent cell. The haematopoietic progenitor cell cultured using stem cell growth factor-like proteins can replace as a graft for the bone marrow transplantation or cord blood transplantation for treating a variety of diseases such as immunodeficiency syndrome, chronic granulomatous disease, duplicated immunodeficiency syndrome, agammaglobulinaemia, Wiskott-Aldrich syndrome, acquired immune deficiency syndrome (AIDS), thalassaemia, haemolytic anaemia due to enzyme defect, congenital anaemia such as sickle cell anaemia, Gaucher's disease, lysosomal storage diseases such as mucopolysaccharidoses, adrenoleukodystrophy, degeneration, a variety of cancer and tumours. Proteins of the invention are useful for treating diseases such as Parkinson's disease, Alzheimer's disease and other neurodegenerative diseases, thrombocytopaenia, immune deficiencies and disorders such as severe combined immunodeficiency (SCID) and autoimmune disorders such as multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, and autoimmune pulmonary inflammation. Sequences of the invention are also useful in gene therapy. The present sequence is human SCR-1 related protein.

XX SQ Sequence 292 AA;
 XX Query Match 97.1%; Score 1472; DB 22; Length 292;
 Best Local Similarity 100.0%; Pred. No. 1e-108; Mismatches 0; Indels 0; Gaps 0;
 Matches 264; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 HRLISLWLFILNFMEMYIGSONASRRQRMRHMHNVSQCGCATSDYNGLSCKRL 62
 Db 2 HRLISLWLFILNFMEMYIGSONASRRQRMRHMHNVSQCGCATSDYNGLSCKRL 61

Qy 63 FFLALRIGMKQIGVCLSCPSQGYGRYPDINKCTKCKADCTCFNFKNFTCKCSGFYIH 122
 Db 62 FFLALRIGMKQIGVCLSCPSQGYGRYPDINKCTKCKADCTCFNFKNFTCKCSGFYIH 121

QY 123 LGKCLDNCPEGLEANNHTMCVSIVHCEVSEHNPAWSPTKKGKQGFFKFRGTERVRIIQ 182
 CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 122 LGKCLDNCPEGLEANNHTMCVSIVHCEVSEHNPAWSPTKKGKQGFFKFRGTERVRIIQ 181
 CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 183 HPSAKGNLCPPTNETRKCTVORKKQKGKGERKRGKKGKGRKRKPKNGKESKEAIPDSKLESS 241
 CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 182 HPSAKGNLCPPTNETRKCTVORKKQKGKGERKRGKKGKGRKRKPKNGKESKEAIPDSKLESS 241
 CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 243 KEIPORENKOOKKRVQDKKS 266
 CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 242 KEIPORENKOOKKRVQDKKS 265
 CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 RESULT 9
 ABP61846
 ID ABP61846 standard; Protein: 292 AA.
 XX
 AC ABP61846;
 XX
 DT 04-OCT-2002 (first entry)
 XX Human polypeptide SEQ ID NO 200.
 KW Human; cytostatic; antirheumatic; antiarthritic; vulnerary; analgesic;
 KW antinflammatory; antibacterial; antiarthritic; antiparkinsonian;
 KW neuroprotective; notropic; osteopathic; haemostatic; vasotropic;
 KW antiulcer; fungicide; antidiabetic; antiarthritic; antihaemostatic;
 KW immunostimulant; antiparasitic; secreted protein; transmembrane protein;
 KW cytokinin; proliferation; cell differentiation; autoimmune disease;
 KW stem cell; growth factor; nervous system disease; neuropathy;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease;
 KW osteoporosis; severe combined immunodeficiency; SCID; infection;
 KW multiple sclerosis; rheumatoid arthritis; gene therapy.
 OS Homo sapiens.
 XX US2002063394-A1.
 XX PD 30-MAY-2002.
 XX PP 22-DEC-2000; 2000US-0745763.
 PR 18-MAR-1998; 94US-0040963.
 XX
 PA (JACO/.) JACOBS K.
 PA (MCCO/.) MCCOY J M.
 PA (LAV/.) LAVALLIE E R.
 PA (COLL/.) COLLINS-RACIE L A.
 PA (EVAN/.) EVANS C.
 PA (MERB/.) MERBERG D.
 PA (TRE/.) TREACY M.
 PA (SPAU/.) SPAULDING V.
 XX
 PT Jacobs K, MCCOY JM, Lavallie ER, Collins-Racie LA, Evans C;
 PT Merberg D, Treacy M, Spaulding V;
 XX WPI; 2002-582343/62.
 DR -PSDB; ABQ92060.
 XX
 PT Novel secreted or transmembrane protein and polynucleotide encoding the
 PT protein, useful for diagnosis and treatment of neurological disorders,
 PT cancer, autoimmune diseases, bone disorders and lung or liver fibrosis
 PT
 XX
 PS Claim 207; Page 203-204; 284pp; English.
 XX
 CC The invention relates to human secreted or transmembrane protein (I),
 CC their fragments and is encoded by specific complementary deoxyribonucleic
 acid (cDNA) inserts (II), where the protein is substantially free from
 other mammalian proteins. (I) are useful for preventing, treating or
 ameliorating a medical condition, especially immunological treatment or
 prevention of tumours. (I) exhibits activity relating to angiogenesis,

CC cytokine, cell proliferation, cell differentiation, antiinflammatory,
 CC stem cell, growth factor activity and activin or inhibin-related
 CC activities. (I) can be used to manipulate stem cells in culture to give
 rise to neuroepithelial cells that can be used to augment or replace
 cells damaged by illness, autoimmune disease, accidental damage or
 genetic disorders. (I) induces the proliferation of neural cells and
 regeneration of nerve and brain tissue and is useful for the treatment of
 central and peripheral nervous system diseases and neuropathies, such as
 Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 lateral sclerosis. (I) is involved in chemotactic or chemokinetic
 activity, regulation of haemopoiesis and is useful for treating myeloid
 or lymphoid cell disorders, platelet disorders such as thrombocytopaenia
 and for regeneration of bone, cartilage, tendon, ligament and/or nerve
 tissue growth and in tissue repair, healing of burns, incisions, ulcers,
 CC for treating osteoporosis, osteoarthritis, bone degenerative disorders or
 CC periodontal disease. (I) is also useful for gut protection or
 CC regeneration and treatment of lung or liver fibrosis, reperfusion injury
 CC in various tissues, various immune deficiencies and disorders including
 CC severe combined immunodeficiency (SCID), bacterial or fungal infections,
 CC autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis,
 CC diabetes mellitus, myasthenia gravis, allergic reactions and conditions,
 CC such as asthma or other respiratory problems. (I) is useful to express
 CC recombinant protein, as markers for tissues in which the corresponding
 CC protein is preferentially expressed and in gene therapy. The present
 XX sequence is that of a polypeptide of the invention.

SQ Sequence 292 AA;

Query Match 97.1%; Score 1472; DB 23; Length 292;
 Best Local Similarity 100.0%; Preq. No. 1e-108; 0; Mismatches 0; Gaps 0;
 Matches 264; Conservative 0; Indels 0; Gaps 0;

Db 2 HLRUJSLWLFITIINFMEYIGSQAQRKRRQRRHPPNYSQCGCGATCCTSYNGLSCPKRL 61

QY 63 FFALERGMKQIGVUCSSCSGGYGRTRPDNIKCTKKAQCDTFNPKFCTKKGYHL 122

Db 62 FFALERGMKQIGVUCSSCSGGYGRTRPDNIKCTKKAQCDTFNPKFCTKKGYHL 121

QY 123 LGKCLDNCPEGLEANNHTMCVSIVHCEVSEHNPAWSPTKKGKQGFFKFRGTERVRIIQ 182
 CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 122 LGKCLDNCPEGLEANNHTMCVSIVHCEVSEHNPAWSPTKKGKQGFFKFRGTERVRIIQ 181
 CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 183 HPSAKGNLCPPTNETRKCTVORKKQKGKGERKRGKKGKGRKRKPKNGKESKEAIPDSKLESS 241
 CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 182 HPSAKGNLCPPTNETRKCTVORKKQKGKGERKRGKKGKGRKRKPKNGKESKEAIPDSKLESS 241
 CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 243 KEIPORENKOOKKRVQDKKS 266
 CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 242 KEIPORENKOOKKRVQDKKS 265
 CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 RESULT 10
 ID AAB93875 standard; Protein: 292 AA.
 XX
 AC AAB93875;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:13781.
 XX
 KW 'Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PP 28-JUL-2000; 2000EP-0116126.

PR 29-JUL-1999; 990JP-0248036.
 PR 27-AUG-1999; 990JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyma T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs.
 XX
 PS Claim 8; SEQ ID 13781; 2537pp + CD ROM; English.
 XX
 CC The present invention describes primer sets for synthesising 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH0166 to AAH1628 and
 CC AAH9593 represent human amino acid sequences; and AAH3629 to AAH3632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 SQ Sequence 292 AA;
 CC
 Query Match 95.8%; Score 1452; DB 22; Length 292;
 CC
 Best Local Similarity 99.2%; Pred. No. 3; 8e-107; Length 292;
 CC
 Matches 262; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 CC
 OY 3 HRLRLISWLFILNPEMEXYGSQNSASRRQRMRHMNSQCGCGACGCSDPNGCLSKRL 62
 DB 2 HRLRLISWLFILNPEMEXYGSQNSASRRQRMRHMNSQCGCGACGCSDPNGCLSKRL 61
 OY 63 FPLALRIGMIGQIGVCLSCPSCPYYGGRYRPIKTKRADCTCFKINFCFKCKSGFYHL 122
 DB 62 FPLALRIGMIGQIGVCLSCPSCPYYGGRYRPIKTKRADCTCFKINFCFKCKSGFYHL 121
 OY 123 LGKCDUDNCERGLPANNHMECVSIVVCESEWNWPSCTKKTCKPGRMTRERIQL 182
 DB 122 LGKCDUDNCERGLPANNHMECVSIVVCESEWNWPSCTKKTCKPGRMTRERIQL 181
 OY 183 HPSAKGNLCPPTETRKCTVORKKCKGKGKRRERKKGKGSKEAPDSKLISS 242
 DB 182 HPSAKGNLCPPTETRKCTVORKKCKGKGKRRERKKGKGSKEAPDSKLISS 241
 OY 243 KEIPPORENKQOKKRKVQDKKS 266
 DB 242 KEIPPORENQOKKRKVQDKKS 265
 RESULT 11
 AAI1313 standard; Protein: 251 AA.
 ID AAE13153

XX
 AC AAE13153;
 XX
 DT 28-JAN-2002 (first entry)
 XX
 DE Human mature stem cell growth factor-like protein.
 XX
 KW Human; stem cell growth factor-like protein; antiinflammatory; nocropic;
 KW neuroprotective; vulnerary; cytostatic; anticonvulsant; immunostimulant;
 KW vasotropic; virucide; dermatological; tranquilliser; cerebroprotective;
 KW osteopathic; immunodeficiency syndrome; chronic granulomatous disease;
 KW duplicated immune deficiency syndrome; Wiskott-Aldrich syndrome; AIDS;
 KW acquired immune deficiency syndrome; agammaglobulinaemia; thalassemia;
 KW Gaucher's disease; lysosomal storage disease; mucopolysaccharidosis;
 KW adrenal white matter degeneration; anaemia; neurodegenerative disease;
 KW Parkinson's disease; Alzheimer's disease; thrombocytopaenia; SCID;
 KW severe combined immunodeficiency; immune disorder; autoimmune disorder;
 KW multiple sclerosis; systemic lupus erythematosus; rheumatoid arthritis;
 KW autoimmune pulmonary inflammation; cancer; tumour; gene therapy;
 KW SCR-1; supporting factor for the proliferation of stem cell.
 XX
 OS Homo sapiens.
 XX
 PN WO200117169-A2.
 XX
 PD 18-OCT-2001.
 XX
 PF 05-APR-2001; 2001WO-US11208.
 XX
 PR 05-APR-2000; 2000US-0543774.
 PR 28-JUN-2000; 2000US-2157339.
 PR 09-JAN-2001; 2001US-0757532.
 PR 05-FEB-2001; 2001US-266614P.
 PR PA (HYSE-) HYSEQ INC.
 PA (KIRI) KIRIN BEER KK.
 XX
 PI Tang TY, Labat I, Tillinghast JS, Sinku A, Liu C, Drmanac RT;
 PI Stache-Crain B, Dickson M, Mize NK, Nishikawa M;
 XX
 DR WPI; 2001-657166/75.
 XX
 PT Novel stem cell growth factor like polypeptides and polynucleotides for
 PT identifying modulators useful for treating diseases such as Alzheimer's
 PT disease, cancer, rheumatoid arthritis, osteoporosis
 XX
 PS Claim 28; Page 216-217; 232pp; English.
 CC The patient discloses novel stem cell growth factor-like proteins and
 CC polynucleotides encoding them. Proteins of the invention are also known
 CC as supporting factor for the proliferation of stem cells (SCR-1). Stem
 CC cell growth factor-like proteins are useful for supporting proliferation
 CC of a stem cell or germ cell which is preferably primordial
 CC germ cell, germ line stem cell, embryonic stem cell, haematopoietic stem
 CC cell, haematopoietic progenitor cell, pluripotent cell or totipotent cell.
 CC The haematopoietic progenitor cell culture using stem cell
 CC growth factor-like proteins can replace as a graft for the bone marrow
 CC transplantation or cord blood transplantation for treating a variety
 CC of diseases such as immunodeficiency syndrome, chronic granulomatous
 CC disease, duplicated immunodeficiency syndrome, agammaglobulinaemia,
 CC Wiskott-Aldrich syndrome, acquired immune deficiency syndrome (AIDS),
 CC thalassaemia, haemolytic anaemia due to enzyme defect, congenital
 CC anaemia such as sickle cell anaemia, Gaucher's disease, lysosomal
 CC storage diseases such as mucopolysaccharidoses, adrenal white matter
 CC degeneration, a variety of cancer and tumours. Proteins of the
 CC invention are useful for treating diseases such as Parkinson's
 CC disease, Alzheimer's disease and other neurodegenerative diseases,
 CC thrombocytopaenia, immune deficiencies and disorders such as severe
 CC combined immunodeficiency (SCID) and autoimmune disorders such as
 CC multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis,
 CC and autoimmune pulmonary inflammation. Sequences of the invention are
 CC also useful in gene therapy. The present sequences is the mature protein
 CC of human stem cell growth factor-like protein.

CC polypeptides of the invention have homology to known proteins, thereby giving an insight into their probable biological activities, and hence potential therapeutic applications. The polypeptides of the invention may have various activities, including cytokine, cell proliferation or cell differentiation activities; stem cell growth factor activity; haemopoiesis regulatory activity; tissue growth factor activity; immunomodulatory activity; activin- or inhibin-related activities; chemotactic or chemokinetic activities; haemostatic, thrombotic or thrombolytic activities; receptor or ligand activities; or may be involved in oncogenesis, cancer cell proliferation or metastasis.

CC Depending on their biological activities, polypeptides and nucleotides of the invention are useful for preventing, treating or ameliorating medical conditions, e.g., by protein or gene therapy. Such conditions include cancers, haemopoietic disorders (e.g., myeloid or lymphoid cell disorders), chronic inflammatory conditions (e.g., asthma or arthritis), proliferative retinopathy, atherosclerosis, coronary heart disease, arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal vascular growth. Polypeptides involved with tissue regeneration and repair (or nucleic acids encoding them) may be used to promote wound healing (e.g., of burns, incisions and ulcers), while those with immunomodulatory activities may be used in the treatment of viral, bacterial and fungal infections in addition to immune disorders.

CC Polypeptides with growth factor activity may be used in cell cultures to promote cell growth. For example, such polypeptides may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease or accidental damage. The polypeptides and nucleotides may also be used in the diagnosis of the above conditions, and in drug screening techniques. The present sequence represents a novel human polypeptide of the invention.

CC SQ Sequence 160 AA;

Query Match Similarity 59.6%; Score 903; DB 22; Length 160; Best Local Similarity 100.0%; Pred. No. 5.8e-64; Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 CTKCKADCDTCFNMNFKPCKCKSGFPVYHLGKCLDNCPGELEANHTMCEVSVI VHCEVSEVN 155

Db 1 CTKCKADCDTCFNMNFKPCKCKSGFPVYHLGKCLDNCPGELEANHTMCEVSVI VHCEVSEVN 60

QY 156 PWSPTKKGKTCFGKRGTETRVREIIOHPSAKGNLCPPTNETRKCTVORKKCQKGERGK 215

Db 61 PWSPTKKGKTCFGKRGTETRVREIIOHPSAKGNLCPPTNETRKCTVORKKCQKGERGK 120

QY 216 GRERKPKNKGESKEAIPDSKLSSKEIPEORENKKQ 255

Db 121 GRERKPKNKGESKEAIPDSKLSSKEIPEORENKKQ 160

RESULT 15

AAW79312

ID AAW79312 standard; Protein: 160 AA.

AC AAW79312;

XX

DT 06-NOV-2001 (first entry)

XX Human protein SEQ ID NO 2958.

DE

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy; vaccine; peptide therapy; stem cell growth factor; haemopoiesis; tissue growth factor; immunomodulatory; cancer; leukaemia; nervous system disorder; arthritis; inflammation.

XX Homo sapiens.

OS Homo sapiens.

XX WO200157190-A2.

XX PD 09-AUG-2001.

XX

XX 05-FEB-2001; 2001WO-US04098.

PR 03-FEB-2000; 2000US-0496914.

PR 27-APR-2000; 2000US-0560875.

PR 20-JUN-2000; 2000US-0598075.

PR 19-JUL-2000; 2000US-0620325.

PR 01-SEP-2000; 2000US-0654936.

PR 15-SEP-2000; 2000US-0633561.

PR 20-OCT-2000; 2000US-0793135.

PR 30-NOV-2000; 2000US-0728422.

(HYSE-) HYSEQ INC.

PA

XX

PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;

XX PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

PI Xie AJ, Yang Y, Weijerman T, Goodrich R;

XX DR WPI; 2001-476283/51.

DR N-PSDB; AAK52445.

XX PT Nucleic acids encoding polypeptides with cytokine-like activities, useful in diagnosis

XX PS Claim 20; Page 214; 621pp; English.

CC The invention relates to polynucleotides (AAK51456-AAK53435) and the encoded polypeptides (AAW8322-AAW80302) that exhibit activity elating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The

CC polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haemopoiesis regulating

CC activity, tissue growth factor activity, immunomodulatory activity and activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation.

CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666 (AAW80201) are omitted as the relevant pages from the sequence listing were missing at the time of publication.

XX SQ Sequence 160 AA;

Query Match

Best Local Similarity 59.6%; Score 903; DB 22; Length 160; Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 CTKCKADCDTCFNMNFKPCKCKSGFPVYHLGKCLDNCPGELEANHTMCEVSVI VHCEVSEVN 155

Db 1 CTKCKADCDTCFNMNFKPCKCKSGFPVYHLGKCLDNCPGELEANHTMCEVSVI VHCEVSEVN 60

QY 156 PWSPTKKGKTCFGKRGTETRVREIIOHPSAKGNLCPPTNETRKCTVORKKCQKGERGK 215

Db 61 PWSPTKKGKTCFGKRGTETRVREIIOHPSAKGNLCPPTNETRKCTVORKKCQKGERGK 120

QY 216 GRERKPKNKGESKEAIPDSKLSSKEIPEORENKKQ 255

Db 121 GRERKPKNKGESKEAIPDSKLSSKEIPEORENKKQ 160

RESULT 15

AAW79312

ID AAW79312 standard; Protein: 160 AA.

AC AAW79312;

XX

DT 06-NOV-2001 (first entry)

XX Human protein SEQ ID NO 2958.

DE

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;

XX vaccine; peptide therapy; stem cell growth factor; haemopoiesis;

XX tissue growth factor; immunomodulatory; cancer; leukaemia;

XX nervous system disorder; arthritis; inflammation.

XX

OS Homo sapiens.

XX WO200157190-A2.

XX PD 09-AUG-2001.

XX

XX 05-FEB-2001; 2001WO-US04098.

Search completed: May 6, 2003, 14:51:39

Job time : 38 secs

PO55_BRACL PCK5_BRACL STANDARD PRT; 1696 AA.
 ID ID
 AC 09N15; 09N16; 09N14;
 DT DT
 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 0, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE DE
 Proprotein convertase subtilisin/kexin type 5 precursor (EC 3.4.21.-)
 DE (Proprotein convertase PC6-like) (PC6).
 GN GN
 PC6.
 OS Branchiostoma californiensis (California lancelet) (Amphioxus).
 OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomida;
 OC Branchiostoma
 NCBI_TaxID=7738;
 OX OX
 RN [1]--> FROM N.A. [(ISOFORMS A, B AND C).
 RP SEQUENCE FROM N.A. [(ISOFORMS A, B AND C).
 RX MEDLINE=0172281; PubMed=1070866;
 RA Oliva A. A. JR.; Chan S. J.; Steiner D. F.;
 RT "Evolution of the prohormone convertases: identification of a
 homologue of PC6 in the protochordate amphioxus.";
 RL Biochim. Biophys. Acta 1477:338-348 (2000).
 CC -- FUNCTION: LIKELY TO REPRESENT A WIDESPREAD ENDOPROTEASE ACTIVITY
 CC WITHIN THE CONSTITUTIVE AND REGULATED SECRETORY PATHWAY CAPABLE
 CC OF CLEAVAGE AT THE RX(K/R)R CONSENSUS MOTIF (BY SIMILARITY).
 CC -- CATALYTIC ACTIVITY: RELEASE OF MATURE PROTEINS FROM THEIR
 CC PROTEONES BY CLEAVAGE OF ARG-XAA-YAA-ARG| -ZAA BONDS, WHERE XAA
 CC CAN BE ANY AMINO ACID AND YAA IS ARG OR LYS.
 CC -- SUBCELLULAR LOCATION: ISOFORM A AND ISOFORM C ARE SECRETED.
 CC ISOFORM B IS A TYPE I MEMBRANE PROTEIN.
 CC -- ALTERNATIVE PRODUCTS: 3 ISOFORMS; A, B (SHOWN HERE) AND C; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -- DOMAIN: THE PROPEPTIDE DOMAIN ACTS AS AN INTRAMOLECULAR CHAPERONE
 CC ASSISTING THE FOLDING OF THE ZYMOGEN WITHIN THE ENDOPLASMIC
 CC RETICULUM.
 CC -- SIMILARITY: BELONGS TO PEPTIDASE FAMILY SB.
 CC -- SIMILARITY: CONTAINS 1 HOMO BTP DOMAIN.
 CC
 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC
 DR EMBL; AFR84615; AAF26300.1; --.
 DR EMBL; AFR84617; AAF26302.1; --.
 DR EMBL; AFR84617; AAF26302.1; --.
 DR HSSP; Q98405; IMPR.
 DR MEROPS; S08_UPB; --.
 DR InterPro; IPR00561; EGF-like.
 DR InterPro; IPR002174; Furin-like.
 DR InterPro; IPR002884; P-domain.
 DR InterPro; IPR00209; Peptidase_S8.
 DR Pfam; PRO008; Peptidase_S8; 1.
 DR PRINTS; PRO00723; SUBILISIN.
 DR PRODOM; P0000717; P-domain, 1.
 DR SMART; S000181; EGF; 2.
 DR SMART; S00261; FU; 17.
 DR PROSITE; PS00136; SUBILASE ASP; FALSE_NEG.
 DR PROSITE; PS00117; SUBILASE HIS; 1.
 DR PROSITE; PS00138; SUBILASE SER; 1.
 KW Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal;
 KW Cleavage on pair of basic residues; Repeat; Alternative splicing;
 KW Transmembrane.
 FT SIGNAL 1 25 POTENTIAL.
 FT PROPEP 26 110 POTENTIAL.
 FT CHAIN 111 1696 PROTEIN CONVERTASE SUBILISIN/KEXIN
 FT DOMAIN 111 1619 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1619 1619 POTENTIAL.
 FT DOMAIN 1640 1640 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 111 488 CATALYTIC (POTENTIAL).
 RN [1]--> FROM N.A.
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9238103; PubMed=1512259;
 RA Rebroek, A.J.M.; Creemers, J.W.M.; Pauli, I.G.L.; Kurzik-Dumke, U.;
 RA Renkrop, M.; Gateroff, E.A.F.; Leunissen, J.A.M.; van de Ven, W.J.;
 RA "Cloning and functional expression of Dtrun1, a subtilisin-like
 RT protease processing enzyme of *Drosophila melanogaster* with multiple
 RT repeats of a cysteine motif";
 RL J. Biol. Chem. 267:17205-17215 (1992).
 CC -- FUNCTION: FURIN IS LIKELY TO REPRESENT THE UNIQUitous ENDOPROTEASE
 CC ACTIVITY WITHIN CONSTITUTIVE SECRETORY PATHWAYS AND CAPABLE OF
 CC CLEAVAGE AT THE RX(K/R)R CONSENSUS MOTIF (BY SIMILARITY).
 CC -- CATALYTIC ACTIVITY: Release of mature proteins from their

proproteins by cleavage of Arg-Xaa-Yaa-Arg-|-Zaa bonds, where Xaa can be any amino acid and Yaa is Arg or Lys. Releases albumin, complement component C3 and von Willebrand factor from their respective precursors.

!-! SIMILARITY: BELONGS TO PEPTIDASE FAMILY SB: FURIN SUBFAMILY.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC

CC EMBL: M94375; AAA28551.1; -.

CC PIR: A43334; A43334.

CC HSSP: Q99405; IMPR.

CC MEROPS: S08.049; -.

CC FlyBase: FBgn000598; Fur2.

CC InterPro: IPR000561; EGFR-like.

CC InterPro: IPR002174; Furin-like.

CC InterPro: IPR002841; P-domain.

CC InterPro: IPR000209; Peptidase_S8.

CC Pfam: PF00082; Peptidase_S8; 1.

CC Pfam: PF0183; P; PRMTL.

CC PRINTS: PR00723; SUBTILISIN.

CC PRODOM: PD000717; P-domain; 1.

CC SMART: SM0261; FU; 10.

CC PROSITE: PS00136; SUBTILASE_ASPP; 1.

CC PROSITE: PS00137; SUBTILASE_HSS; 1.

CC PROSITE: PS00138; SUBTILASE_SRR; 1.

CC Hydrolase; Serine Protease; Glycoprotein; Signal; Transmembrane; Multigene family; Zymogen; Repeat; SIGNAL; ?; POTENTIAL.

CC PROPEP: ?; 319.

CC CHAIN: 320 1680 PURIN-LIKE PROTEASE 2.

CC ACT SITE: 418 418 CHARGE RELAY SYSTEM (BY SIMILARITY).

CC ACT SITE: 457 457 CHARGE RELAY SYSTEM (BY SIMILARITY).

CC ACT SITE: 638 638 CHARGE RELAY SYSTEM (BY SIMILARITY).

CC DOMAIN: 962 1444 10 X TANDEM REPEATS, CYS-RICH.

CC REPEAT: 962 1007 1.

CC REPEAT: 1008 1057 2.

CC REPEAT: 1058 1104 3.

CC REPEAT: 1105 1153 4.

CC REPEAT: 1154 1205 5.

CC REPEAT: 1206 1254 6.

CC REPEAT: 1255 1299 7.

CC REPEAT: 1300 1346 8.

CC REPEAT: 1347 1393 9.

CC REPEAT: 1394 1444 10.

CC TRANSMEM: 1508 1532 POTENTIAL.

CC DOMAIN: 1533 1680 CYTOSOLIC (POTENTIAL).

CC CARBOHYD: 3 3 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD: 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD: 130 130 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD: 203 203 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD: 443 443 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD: 481 481 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD: 928 928 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD: 1061 1061 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD: 1182 1182 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD: 1275 1275 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD: 1278 1278 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD: 1440 1440 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC SEQUENCE: 1680 AA; 183599 MN; 0A99CE8770A88293 CRC64;

Query Match Score: 11.6%; Score: 176.5; DB: 1; Length: 1680; Best Local Similarity: 28.0%; Preq. No. 3; 3e-05; Matches: 60; Conservative: 24; Mismatches: 77; Indels: 53; Gaps: 12; Qy 3 HRLRLISWLFILNPMEVIGSONASGRQRORRMPNVSQGCGCTCSDY-NGCLSLKPR 61

Db 1038 HLHVID-LAVQHQCPDGFENS---RNRCTVPP---CEPNCAASCOPHEPEYTSCHH 1087

Qy 62 LFFALERIGMKGQIVLSSCPGSGYTRPDRDINKCTCKKACDDCF--NKNFTCKCSGF 119

Db 1088 LYMEBHK-----GYSACPDLYET--EDNKCFCFHSCATONGPPDQCITCNSR 1136

Qy 120 YLHIGKCLDNCPBGLAANNHMECVSIHCESEWWMSPCKKGTGPKRGTRTRY-R 178

Db 1137 YAWQNKCLISCDPGFVADKRCLEW-----PCQEGKTC---TSNGVCS 1177

Qy 179 EIQQPSAKGNLCPPTNETRKCTWKK-CQKGB 211

Db 1178 ECLQWNT-----LNKRDKCIVSGSGCSE 1203

RESULT 4

PAC4_HUMAN STANDARD; PRT; 969 AA.

ID PAC4_HUMAN STANDARD; PRT; 969 AA.

ID PAC4_HUMAN STANDARD; PRT; 969 AA.

AC P29112; Q15059; Q15100; Q9UEJ1; Q9UEJ2; Q9UEJ3; Q9UEJ9;

AC Q9UEJ7; Q9Y4G9; Q9Y4H1; Q9Y4H2; Q9Y4H3; Q9Y4H4; Q9Y4H5;

AC Q9Y4H6; Q9Y4H7; Q9Y4H8; Q9Y4H9; Q9Y4H10; Q9Y4H11; Q9Y4H12;

AC Q9Y4H13; Q9Y4H14; Q9Y4H15; Q9Y4H16; Q9Y4H17; Q9Y4H18; Q9Y4H19;

AC Q9Y4H20; Q9Y4H21; Q9Y4H22; Q9Y4H23; Q9Y4H24; Q9Y4H25; Q9Y4H26;

AC Q9Y4H27; Q9Y4H28; Q9Y4H29; Q9Y4H30; Q9Y4H31; Q9Y4H32; Q9Y4H33;

AC Q9Y4H34; Q9Y4H35; Q9Y4H36; Q9Y4H37; Q9Y4H38; Q9Y4H39; Q9Y4H40;

AC Q9Y4H41; Q9Y4H42; Q9Y4H43; Q9Y4H44; Q9Y4H45; Q9Y4H46; Q9Y4H47;

AC Q9Y4H48; Q9Y4H49; Q9Y4H50; Q9Y4H51; Q9Y4H52; Q9Y4H53; Q9Y4H54;

AC Q9Y4H55; Q9Y4H56; Q9Y4H57; Q9Y4H58; Q9Y4H59; Q9Y4H60; Q9Y4H61;

AC Q9Y4H62; Q9Y4H63; Q9Y4H64; Q9Y4H65; Q9Y4H66; Q9Y4H67; Q9Y4H68;

AC Q9Y4H69; Q9Y4H70; Q9Y4H71; Q9Y4H72; Q9Y4H73; Q9Y4H74; Q9Y4H75;

AC Q9Y4H76; Q9Y4H77; Q9Y4H78; Q9Y4H79; Q9Y4H80; Q9Y4H81; Q9Y4H82;

AC Q9Y4H83; Q9Y4H84; Q9Y4H85; Q9Y4H86; Q9Y4H87; Q9Y4H88; Q9Y4H89;

AC Q9Y4H90; Q9Y4H91; Q9Y4H92; Q9Y4H93; Q9Y4H94; Q9Y4H95; Q9Y4H96;

AC Q9Y4H97; Q9Y4H98; Q9Y4H99; Q9Y4H100; Q9Y4H101; Q9Y4H102; Q9Y4H103;

AC Q9Y4H104; Q9Y4H105; Q9Y4H106; Q9Y4H107; Q9Y4H108; Q9Y4H109; Q9Y4H110;

AC Q9Y4H111; Q9Y4H112; Q9Y4H113; Q9Y4H114; Q9Y4H115; Q9Y4H116; Q9Y4H117;

AC Q9Y4H118; Q9Y4H119; Q9Y4H120; Q9Y4H121; Q9Y4H122; Q9Y4H123; Q9Y4H124;

AC Q9Y4H125; Q9Y4H126; Q9Y4H127; Q9Y4H128; Q9Y4H129; Q9Y4H130; Q9Y4H131;

AC Q9Y4H132; Q9Y4H133; Q9Y4H134; Q9Y4H135; Q9Y4H136; Q9Y4H137; Q9Y4H138;

AC Q9Y4H139; Q9Y4H140; Q9Y4H141; Q9Y4H142; Q9Y4H143; Q9Y4H144; Q9Y4H145;

AC Q9Y4H146; Q9Y4H147; Q9Y4H148; Q9Y4H149; Q9Y4H150; Q9Y4H151; Q9Y4H152;

AC Q9Y4H153; Q9Y4H154; Q9Y4H155; Q9Y4H156; Q9Y4H157; Q9Y4H158; Q9Y4H159;

AC Q9Y4H160; Q9Y4H161; Q9Y4H162; Q9Y4H163; Q9Y4H164; Q9Y4H165; Q9Y4H166;

AC Q9Y4H167; Q9Y4H168; Q9Y4H169; Q9Y4H170; Q9Y4H171; Q9Y4H172; Q9Y4H173;

AC Q9Y4H174; Q9Y4H175; Q9Y4H176; Q9Y4H177; Q9Y4H178; Q9Y4H179; Q9Y4H180;

AC Q9Y4H181; Q9Y4H182; Q9Y4H183; Q9Y4H184; Q9Y4H185; Q9Y4H186; Q9Y4H187;

AC Q9Y4H188; Q9Y4H189; Q9Y4H190; Q9Y4H191; Q9Y4H192; Q9Y4H193; Q9Y4H194;

AC Q9Y4H195; Q9Y4H196; Q9Y4H197; Q9Y4H198; Q9Y4H199; Q9Y4H200; Q9Y4H201;

AC Q9Y4H202; Q9Y4H203; Q9Y4H204; Q9Y4H205; Q9Y4H206; Q9Y4H207; Q9Y4H208;

AC Q9Y4H209; Q9Y4H210; Q9Y4H211; Q9Y4H212; Q9Y4H213; Q9Y4H214; Q9Y4H215;

AC Q9Y4H216; Q9Y4H217; Q9Y4H218; Q9Y4H219; Q9Y4H220; Q9Y4H221; Q9Y4H222;

AC Q9Y4H223; Q9Y4H224; Q9Y4H225; Q9Y4H226; Q9Y4H227; Q9Y4H228; Q9Y4H229;

AC Q9Y4H230; Q9Y4H231; Q9Y4H232; Q9Y4H233; Q9Y4H234; Q9Y4H235; Q9Y4H236;

AC Q9Y4H237; Q9Y4H238; Q9Y4H239; Q9Y4H240; Q9Y4H241; Q9Y4H242; Q9Y4H243;

AC Q9Y4H244; Q9Y4H245; Q9Y4H246; Q9Y4H247; Q9Y4H248; Q9Y4H249; Q9Y4H250;

AC Q9Y4H251; Q9Y4H252; Q9Y4H253; Q9Y4H254; Q9Y4H255; Q9Y4H256; Q9Y4H257;

AC Q9Y4H258; Q9Y4H259; Q9Y4H260; Q9Y4H261; Q9Y4H262; Q9Y4H263; Q9Y4H264;

AC Q9Y4H265; Q9Y4H266; Q9Y4H267; Q9Y4H268; Q9Y4H269; Q9Y4H270; Q9Y4H271;

AC Q9Y4H272; Q9Y4H273; Q9Y4H274; Q9Y4H275; Q9Y4H276; Q9Y4H277; Q9Y4H278;

AC Q9Y4H279; Q9Y4H280; Q9Y4H281; Q9Y4H282; Q9Y4H283; Q9Y4H284; Q9Y4H285;

AC Q9Y4H286; Q9Y4H287; Q9Y4H288; Q9Y4H289; Q9Y4H290; Q9Y4H291; Q9Y4H292;

AC Q9Y4H293; Q9Y4H294; Q9Y4H295; Q9Y4H296; Q9Y4H297; Q9Y4H298; Q9Y4H299;

AC Q9Y4H299; Q9Y4H300; Q9Y4H301; Q9Y4H302; Q9Y4H303; Q9Y4H304; Q9Y4H305;

AC Q9Y4H306; Q9Y4H307; Q9Y4H308; Q9Y4H309; Q9Y4H310; Q9Y4H311; Q9Y4H312;

AC Q9Y4H313; Q9Y4H314; Q9Y4H315; Q9Y4H316; Q9Y4H317; Q9Y4H318; Q9Y4H319;

AC Q9Y4H320; Q9Y4H321; Q9Y4H322; Q9Y4H323; Q9Y4H324; Q9Y4H325; Q9Y4H326;

AC Q9Y4H327; Q9Y4H328; Q9Y4H329; Q9Y4H330; Q9Y4H331; Q9Y4H332; Q9Y4H333;

AC Q9Y4H334; Q9Y4H335; Q9Y4H336; Q9Y4H337; Q9Y4H338; Q9Y4H339; Q9Y4H340;

AC Q9Y4H341; Q9Y4H342; Q9Y4H343; Q9Y4H344; Q9Y4H345; Q9Y4H346; Q9Y4H347;

AC Q9Y4H348; Q9Y4H349; Q9Y4H350; Q9Y4H351; Q9Y4H352; Q9Y4H353; Q9Y4H354;

AC Q9Y4H355; Q9Y4H356; Q9Y4H357; Q9Y4H358; Q9Y4H359; Q9Y4H360; Q9Y4H361;

AC Q9Y4H362; Q9Y4H363; Q9Y4H364; Q9Y4H365; Q9Y4H366; Q9Y4H367; Q9Y4H368;

AC Q9Y4H369; Q9Y4H370; Q9Y4H371; Q9Y4H372; Q9Y4H373; Q9Y4H374; Q9Y4H375;

AC Q9Y4H376; Q9Y4H377; Q9Y4H378; Q9Y4H379; Q9Y4H380; Q9Y4H381; Q9Y4H382;

AC Q9Y4H383; Q9Y4H384; Q9Y4H385; Q9Y4H386; Q9Y4H387; Q9Y4H388; Q9Y4H389;

AC Q9Y4H390; Q9Y4H391; Q9Y4H392; Q9Y4H393; Q9Y4H394; Q9Y4H395; Q9Y4H396;

AC Q9Y4H397; Q9Y4H398; Q9Y4H399; Q9Y4H400; Q9Y4H401; Q9Y4H402; Q9Y4H403;

AC Q9Y4H404; Q9Y4H405; Q9Y4H406; Q9Y4H407; Q9Y4H408; Q9Y4H409; Q9Y4H410;

AC Q9Y4H411; Q9Y4H412; Q9Y4H413; Q9Y4H414; Q9Y4H415; Q9Y4H416; Q9Y4H417;

AC Q9Y4H418; Q9Y4H419; Q9Y4H420; Q9Y4H421; Q9Y4H422; Q9Y4H423; Q9Y4H424;

AC Q9Y4H425; Q9Y4H426; Q9Y4H427; Q9Y4H428; Q9Y4H429; Q9Y4H430; Q9Y4H431;

AC Q9Y4H432; Q9Y4H433; Q9Y4H434; Q9Y4H435; Q9Y4H436; Q9Y4H437; Q9Y4H438;

AC Q9Y4H439; Q9Y4H440; Q9Y4H441; Q9Y4H442; Q9Y4H443; Q9Y4H444; Q9Y4H445;

AC Q9Y4H446; Q9Y4H447; Q9Y4H448; Q9Y4H449; Q9Y4H450; Q9Y4H451; Q9Y4H452;

AC Q9Y4H453; Q9Y4H454; Q9Y4H455; Q9Y4H456; Q9Y4H457; Q9Y4H458; Q9Y4H459;

AC Q9Y4H460; Q9Y4H461; Q9Y4H462; Q9Y4H463; Q9Y4H464; Q9Y4H465; Q9Y4H466;

AC Q9Y4H467; Q9Y4H468; Q9Y4H469; Q9Y4H470; Q9Y4H471; Q9Y4H472; Q9Y4H473;

AC Q9Y4H474; Q9Y4H475; Q9Y4H476; Q9Y4H477; Q9Y4H478; Q9Y4H479; Q9Y4H480;

AC Q9Y4H481; Q9Y4H482; Q9Y4H483; Q9Y4H484; Q9Y4H485; Q9Y4H486; Q9Y4H487;

AC Q9Y4H488; Q9Y4H489; Q9Y4H490; Q9Y4H491; Q9Y4H492; Q9Y4H493; Q9Y4H494;

AC Q9Y4H495; Q9Y4H496; Q9Y4H497; Q9Y4H498; Q9Y4H499; Q9Y4H500; Q9Y4H501;

AC Q9Y4H502; Q9Y4H503; Q9Y4H504; Q9Y4H505; Q9Y4H506; Q9Y4H507; Q9Y4H508;

AC Q9Y4H509; Q9Y4H510; Q9Y4H511; Q9Y4H512; Q9Y4H513; Q9Y4H514; Q9Y4H515;

AC Q9Y4H516; Q9Y4H517; Q9Y4H518; Q9Y4H519; Q9Y4H520; Q9Y4H521; Q9Y4H522;

AC Q9Y4H523; Q9Y4H524; Q9Y4H525; Q9Y4H526; Q9Y4H527; Q9Y4H528; Q9Y4H529;

AC Q9Y4H530; Q9Y4H531; Q9Y4H532; Q9Y4H533; Q9Y4H534; Q9Y4H535; Q9Y4H536;

AC Q9Y4H537; Q9Y4H538; Q9Y4H539; Q9Y4H540; Q9Y4H541; Q9Y4H542; Q9Y4H543;

AC Q9Y4H544; Q9Y4H545; Q9Y4H546; Q9Y4H547; Q9Y4H548; Q9Y4H549; Q9Y4H550;

AC Q9Y4H551; Q9Y4H552; Q9Y4H553; Q9Y4H554; Q9Y4H555; Q9Y4H556; Q9Y4H557;

AC Q9Y4H558; Q9Y4H559; Q9Y4H560; Q9Y4H561; Q9Y4H562; Q9Y4H563; Q9Y4H564;

AC Q9Y4H565; Q9Y4H566; Q9Y4H567; Q9Y4H568; Q9Y4H569; Q9Y4H570; Q9Y4H571;

AC Q9Y4H572; Q9Y4H573; Q9Y4H574; Q9Y4H575; Q9Y4H576; Q9Y4H577; Q9Y4H578;

AC Q9Y4H579; Q9Y4H580; Q9Y4H581; Q9Y4H582; Q9Y4H583; Q9Y4H584; Q9Y4H585;

AC Q9Y4H586; Q9Y4H587; Q9Y4H588; Q9Y4H589; Q9Y4H590; Q9Y4H591; Q9Y4H592;

AC Q9Y4H593; Q9Y4H594; Q9Y4H595; Q9Y4H596; Q9Y4H597; Q9Y4H598; Q9Y4H599;

AC Q9Y4H599; Q9Y4H600; Q9Y4H601; Q9Y4H602; Q9Y4H603; Q9Y4H604; Q9Y4H605;

AC Q9Y4H606; Q9Y4H607; Q9Y4H608; Q9Y4H609; Q9Y4H610; Q9Y4H611; Q9Y4H612;

AC Q9Y4H613; Q9Y4H614; Q9Y4H615; Q9Y4H616; Q9Y4H617; Q9Y4H618; Q9Y4H619;

AC Q9Y4H620; Q9Y4H621; Q9Y4H622; Q9Y4H623; Q9Y4H624; Q9Y4H625; Q9Y4H626;

AC Q9Y4H627; Q9Y4H628; Q9Y4H629; Q9Y4H630; Q9Y4H631; Q9Y4H632; Q9Y4H633;

AC Q9Y4H634; Q9Y4H635; Q9Y4H636; Q9Y4H637; Q9Y4H638; Q9Y4H639; Q9Y4H640;

AC Q9Y4H641; Q9Y4H642; Q9Y4H643; Q9Y4H644; Q9Y4H645; Q9Y4H646; Q9Y4H647;

AC Q9Y4H648; Q9Y4H649; Q9Y4H650; Q9Y4H651; Q9Y4H652; Q9Y4H653; Q9Y4H654;

AC Q9Y4H655; Q9Y4H656; Q9Y4H657; Q9Y4H658; Q9Y4H659; Q9Y4H660; Q9Y4H661;

AC Q9Y4H662; Q9Y4H663; Q9Y4H664; Q9Y4H665; Q9Y4H666; Q9Y4H667; Q9Y4H668;

AC Q9Y4H669; Q9Y4H670; Q9Y4H671; Q9Y4H672; Q9Y4H673; Q9Y4H674; Q9Y4H675;

AC Q9Y4H676; Q9Y4H677; Q9Y4H678; Q9Y4H679; Q9Y4H680; Q9Y4H681; Q9Y4H682;

AC Q9Y4H683; Q9Y4H684; Q9Y4H685; Q9Y4H686; Q9Y4H687; Q9Y4H688; Q9Y4H689;

AC Q9Y4H689; Q9Y4H690; Q9Y4H691; Q9Y4H692; Q9Y4H693; Q9Y4H694; Q9Y4H695;

AC Q9Y4H696; Q9Y4H697; Q9Y4H698; Q9Y4H699; Q9Y4H700; Q9Y4H701; Q9Y4H702;

AC Q9Y4H703; Q9Y4H704; Q9Y4H705; Q9Y4H706; Q9Y4H707; Q9Y4H708; Q9Y4H709;

AC Q9Y4H709; Q9Y4H710; Q9Y4H711; Q9Y4H712; Q9Y4H713; Q9Y4H714; Q9Y4H715;

AC Q9Y4H716; Q9Y4H717; Q9Y4H718; Q9Y4H719; Q9Y4H720; Q9Y4H721; Q9Y4H722;

AC Q9Y4H723; Q9Y4H724; Q9Y4H725; Q9Y4H726; Q9Y4H727; Q9Y4H728; Q9Y4H729;

AC Q9Y4H729; Q9Y4H730; Q9Y4H731; Q9Y4H732; Q9Y4H733; Q9Y4H734; Q9Y4H735;

AC Q9Y4H736; Q9Y4H737; Q9Y4H738; Q9Y4H739; Q9Y4H740; Q9Y4H741; Q9Y4H742;

AC Q9Y4H743; Q9Y4H744; Q9Y4H745; Q9Y4H746; Q9Y4H747; Q9Y4H748; Q9Y4H749;

AC Q9Y4H749; Q9Y4H750; Q9Y4H751; Q9Y4H752; Q9Y4H753; Q9Y4H754; Q9Y4H755;

AC Q9Y4H756; Q9Y4H757; Q9Y4H758; Q9Y4H759; Q9Y4H760; Q9Y4H761; Q9Y4H762;

AC Q9Y4H763; Q9Y4H764; Q9Y4H765; Q9Y4H766; Q9Y4H767; Q9Y4H768; Q9Y4H769;

AC Q9Y4H769; Q9Y4H770; Q9Y4H771; Q9Y4H772; Q9Y4H773; Q9Y4H774; Q9Y4H775;

AC Q9Y4H775; Q9Y4H776; Q9Y4H777; Q9Y4H778; Q9Y4H779; Q9Y4H780; Q9Y4H781;

AC Q9Y4H781; Q9Y4H782; Q9Y4H783; Q9Y4H784; Q9Y4H785; Q9Y4H786; Q9Y4H787;

AC Q9Y4H787; Q9Y4H788; Q9Y4H789; Q9Y4H790; Q9Y4H791; Q9Y4H792; Q9Y4H793;

AC Q9Y4H793; Q9Y4H794; Q9Y4H795; Q9Y4H796; Q9Y4H797; Q9Y4H798; Q9Y4H799;

AC Q9Y4H799; Q9Y4H800; Q9Y4H801; Q9Y4H802; Q9Y4H803; Q9Y4H804; Q9Y4H805;

AC Q9Y4H805; Q9Y4H806; Q9Y4H807; Q9Y4H808; Q9Y4H809; Q9Y4H810; Q9Y4H811;

AC Q9Y4H811; Q9Y4H812; Q9Y4H813; Q9Y4H814; Q9Y4H815; Q9Y4H816; Q9Y4H817;

AC Q9Y4H817; Q9Y4H818; Q9Y4H819; Q9Y4H820; Q9Y4H821; Q9Y4H822; Q9Y4H823;

AC Q9Y4H823; Q9Y4H824; Q9Y4H825; Q9Y4H826; Q9Y4H827; Q9Y4H828; Q9Y4H829;

AC Q9Y4H829; Q9Y4H830; Q9Y4H831; Q9Y4H832; Q9Y4H833; Q9Y4H834; Q9Y4H835;

AC Q9Y4H835; Q9Y4H836; Q9Y4H837; Q9Y4H838; Q9Y4H839; Q9Y4H840; Q9Y4H841;

AC Q9Y4H841; Q9Y4H842; Q9Y4H843; Q9Y4H844; Q9Y4H845; Q9Y4H846; Q9Y4H847;

AC Q9Y4H847; Q9Y4H848; Q9Y4H849; Q9Y4H850; Q9Y4H851; Q9Y4H852; Q9Y4H853;

AC Q9Y4H853; Q9Y4H854; Q9Y4H855; Q9Y4H856; Q9Y4H857; Q9Y4H858; Q9Y4H859;

AC Q9Y4H859; Q9Y4H860; Q9Y4H861; Q9Y4H862; Q9Y4H863; Q9Y4H864; Q9Y4H865;

AC Q9Y4H865; Q9Y4H866; Q9Y4H867; Q9Y4H868; Q9Y4H869; Q9Y4H870; Q9Y4H871;

AC Q9Y4H871; Q9Y4H872; Q9Y4H873; Q9Y4H874; Q9Y4H875; Q9Y4H876; Q9Y4H877;

AC Q9Y4H877; Q9Y4H878; Q9Y4H879; Q9Y4H880; Q9Y4H881; Q9Y4H882; Q9Y4H883;

AC Q9Y4H883; Q9Y4H884; Q9Y4H885; Q9Y4H886; Q9Y4H887; Q9Y4H888; Q9Y4H889;

AC Q9Y4H889; Q9Y4H890; Q9Y4H891; Q9Y4H892; Q9Y4H893; Q9Y4H894; Q9Y4H895;

AC Q9Y4H895; Q9Y4H896; Q9Y4H897; Q9Y4H898; Q9Y4H899; Q9Y4H900; Q9Y4H901;

AC Q9Y4H901; Q9Y4H902; Q9Y4H903; Q9Y4H904; Q9Y4H905; Q9Y4H906; Q9Y4H907;

AC Q9Y4H907; Q9Y4H908; Q9Y4H909; Q9Y4H910; Q9Y4H911; Q9Y4H912; Q9Y4H913;

AC Q9Y4H913; Q9Y4H914; Q9Y4H915; Q9Y4H916; Q9Y4H917; Q9Y4H918; Q9Y4H919;

AC Q9Y4H919; Q9Y4H920; Q9Y4H921; Q9Y4H

RN [6] SEQUENCE FROM N.A. (ISOFORMS PACE4A-I; A-II; CS; D; E-I; E-II).
 RX SEQUENCE=98021085; Published=9781225;
 RA Tsuji A.; Hine C.; Tamai Y.; Yonemoto K.; Mori K.; Yoshida S.;
 RA Bando M.; Sakai E.; Mori K.; Amanatsu T.; Matsuda Y.;
 RT "Genomic organization and alternative splicing of human PACE4 (SPC4),
 RT kexin-like processing endoprotease.;"
 RL J. Biochem. 122:438-452(1997).
 RN [7] ALTERNATIVE SPLICING (ISOFORM PACE4CS).
 RX MEDLINE=9706424; PubMed=9706861;
 RA Zhong M.; Benjamin S.; Lazure C.; Munzer S.; Seidah N.G.;
 RT "Functional analysis of human PACE4-A and PACE4-C isoforms:
 RT identification of a new PACE4-CS isoform.;"
 RL FEBS Lett. 396:31-36(1996).
 RN [8] CHARACTERIZATION.
 RX MEDLINE=9323359; PubMed=10215603;
 RA Sucic J.F.; Moehring J.M.; Innocencio N.M.; Luchini J.W.;
 RA Moehring T.J.;
 RT "Endoprotease PACE4 is Ca2+-dependent and temperature-sensitive and
 RT can partly rescue the phenotype of a furin-deficient cell strain.;"
 RL Biochem. J. 339:639-647(1999).
 RP
 RX PROCESSING.
 RA MEDLINE=9840849; PubMed=9738469;
 RA Nagahama M.; Taniguchi T.; Hashimoto E.; Inamaki A.; Mori K.;
 RA Tsuji A.; Matuda Y.;
 RT "Biosynthetic processing and quaternary interactions of proprotein
 RT convertase Spc4 (PACE4)."
 RL FEBS Lett. 434:155-159(1998).
 CC -!- FUNCTION: LIKELY TO REPRESENT AN ENDOPROTEASE ACTIVITY WITHIN THE
 CC CONSTITUTIVE SECRETORY PATHWAY, WITH UNIQUE RESTRICTED
 CC DISTRIBUTION IN BOTH NEUROENDOCRINE AND NON-NEUROENDOCRINE TISSUES
 CC AND CAPABLE OF CLEAVAGE AT THE RX (K/R) CONSENSUS MOTIF.
 CC -!- CATALYTIC ACTIVITY: RELEASE OF MATURE PROTEINS FROM THEIR
 CC PROTEOPEPTIDE CLEAVAGE AFFECTS THE RELEASE OF ARG-XAA-YAA-ARG-[ZAA] BONDS,
 CC WHERE XAA CAN BE ANY AMINO ACID AND YAA IS ARG OR LYS.
 CC -!- COFACTOR: PACE4A IS PROBABLY CALCIUM-DEPENDENT.
 CC -!- SUBUNIT: THE PACE4A-I PRECURSOR PROTEIN SEEKS TO EXIST IN THE
 CC RETICULUM ENDOPLASMIC RETICULUM AS BOTH A MONOMER AND A DIMER-SIZED COMPLEX
 CC WHEREAS MATURE PACE4A-I EXISTS ONLY AS A MONOMER, SUGGESTING THAT
 CC PROTEOPEPTIDE CLEAVAGE AFFECTS THE RELEASE OF MATURE PROTEINS FROM THEIR
 CC SUBCELLULAR LOCALIZATION: PACE4A-I AND PACE4-ATT ARE SECRETED. PACE4C
 CC AND PACE4CS ARE NOT SECRETED AND REMAIN PROBABLY IN YOGEN FORM
 IN ENDOPLASMIC RETICULUM. PACE4-I AND PACE4E-I ARE RETAINED
 CC INTRACELLULARLY PROBABLY THROUGH A HYDROPHOBIC CLUSTER IN THEIR C-
 CC TERMINUS. PACE4B MIGHT BE SECRETED.
 CC -!- ALTERNATIVE PRODUCTS: 9 ISOFORMS, PACE4A-I/PACE4 (SHOWN HERE),
 CC PACE4A-II, PACE4B/PACE4-1, PACE4C, PACE4CS, PACE4D, PACE4E-I AND
 CC PACE4E-II, ARE PRODUCED BY ALTERNATIVE SPlicing. ISOFORMS PACE4B,
 CC PACE4E-I AND D MIGHT BE ENZYMICALLY INACTIVE.
 CC -!- TISSUE SPECIFICITY: EACH PACE4 ISOFORM EXHIBITS A UNIQUE
 CC RESTRICTED DISTRIBUTION. PACE4A-I IS EXPRESSED IN HEART, BRAIN,
 CC PLACENTA, LUNG, SKELETAL MUSCLE, KIDNEY, PANCREAS, BUT AT
 CC COMPARATIVELY HIGHER LEVELS IN THE LIVER. PACE4A-II IS AT LEAST
 CC EXPRESSED IN PLACENTA. PACE4 WAS ONLY FOUND IN THE EMBRYONIC
 CC KIDNEY CELL LINE FROM WHICH IT WAS ISOLATED. PACE4C AND PACE4B ARE
 CC EXPRESSED IN PLACENTA. PACE4E-I IS EXPRESSED IN CEREBELLUM,
 CC PLACENTA AND PITUITARY. PACE4E-II IS AT LEAST PRESENT IN
 CC CEREBELLUM.
 CC -!- DOMAIN: THE PROPEPTIDE DOMAIN ACTS AS AN INTRAMOLECULAR CHAPERONE
 CC ASSISTING THE FOLDING OF THE ZYMOGEN WITHIN THE ENDOPLASMIC
 CC RETICULUM. ISOFORM PACE4D LACKS THE PROPEPTIDE DOMAIN.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY SB.
 CC -!- SIMILARITY: CONTAINS 1 HOMO BTP DOMAIN.
 CC
 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to license@ish-sib.ch.
 CC
 CC
 DR M80482; AAAK9981;
 DR EMBL; AB001914; BAA21620.1;
 DR EMBL; AB001898; BAA21620.1; JOINED.
 DR EMBL; AB001900; BAA21620.1; JOINED.
 DR EMBL; AB001901; BAA21620.1; JOINED.
 DR EMBL; AB001902; BAA21620.1; JOINED.
 DR EMBL; AB001903; BAA21620.1; JOINED.
 DR EMBL; AB001904; BAA21620.1; JOINED.
 DR EMBL; AB001905; BAA21620.1; JOINED.
 DR EMBL; AB001914; BAA21621.1; JOINED.
 DR EMBL; AB001898; BAA21621.1; JOINED.
 DR EMBL; AB001900; BAA21621.1; JOINED.
 DR EMBL; AB001901; BAA21621.1; JOINED.
 DR EMBL; AB001902; BAA21621.1; JOINED.
 DR EMBL; AB001903; BAA21621.1; JOINED.
 DR EMBL; AB001904; BAA21621.1; JOINED.
 DR EMBL; AB001905; BAA21621.1; JOINED.
 DR EMBL; AB001906; BAA21621.1; JOINED.
 DR EMBL; AB001907; BAA21622.1; JOINED.
 DR EMBL; AB001908; BAA21622.1; JOINED.
 DR EMBL; AB001909; BAA21621.1; JOINED.
 DR EMBL; AB001910; BAA21621.1; JOINED.
 DR EMBL; AB001911; BAA21622.1; JOINED.
 DR EMBL; AB001912; BAA21622.1; JOINED.
 DR EMBL; AB001913; BAA21622.1; JOINED.
 DR EMBL; AB001901; BAA21622.1; JOINED.
 DR EMBL; AB001902; BAA21622.1; JOINED.
 DR EMBL; AB001903; BAA21622.1; JOINED.
 DR EMBL; AB001904; BAA21622.1; JOINED.
 DR EMBL; AB001905; BAA21622.1; JOINED.
 DR EMBL; AB001906; BAA21622.1; JOINED.
 DR EMBL; AB001907; BAA21624.1; JOINED.
 DR EMBL; AB001908; BAA21623.1; JOINED.
 DR EMBL; AB001909; BAA21623.1; JOINED.
 DR EMBL; AB001910; BAA21623.1; JOINED.
 DR EMBL; AB001911; BAA21623.1; JOINED.
 DR EMBL; AB001912; BAA21624.1; JOINED.
 DR EMBL; AB001913; BAA21624.1; JOINED.
 DR EMBL; AB001905; BAA21625.1; JOINED.
 DR EMBL; AB001906; BAA21625.1; JOINED.
 DR EMBL; AB001907; BAA21625.1; JOINED.
 DR EMBL; AB001908; BAA21625.1; JOINED.
 DR EMBL; AB001909; BAA21625.1; JOINED.
 DR EMBL; AB001910; BAA21625.1; JOINED.

Query Match

11.4%; Score 173.5; DB 1; Length 969;

FT CARBOHYD 804 804 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 854 854 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1710 1710 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1732 1732 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPPLIC 878 915 XXXXXXXXXKXXXXXXKXXXXXXKXXXXXX (IN ISOFORM PS5A)
 FT VARSPPLIC 916 1877 MISSING (IN ISOFORM PS5A).
 SQ 36 PNVHQ-GEQG---GCATCSYNGLISCKRPLFALERIGHQIGVCLSCPSYYGTRY 90
 DB 640 PECBVGCGDPGDHCDFCLAHYKLNTR-----ICVSSCPGHP---H 682
 OY 91 PDINKTCKADPTCNCNKF---CTKCKSGFYIHL-LGKCLDNCPGPLEANNHMECVSI 146
 DB 683 ADKGRKRKCAPNCEESCFCGSHADQCLSKYGFLINEETSCVQACPSYODIKNIC--- 739
 OY 147 VHCEVSEWNPMSCTKKKGTC-GFFRGTETRVREILOHPSAKGNLCPPTNETRKCTV 205
 DB 740 -----GKSENCKTGTGPHNTE-----CKGGL--SLQGSRCSV-- 771
 OY 206 KCGQE 211
 DB 772 TCEBQQ 777

RESULT 6

PCKS_HUMAN STANDARD; PRT; 913 AA.

ID 092874; Q13527; (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Proprotein convertase subtilisin/kexin type 5 precursor (EC 3.4.21.-) (Proprotein convertase PC5) (Subtilisin/kexin-like protease PC5)
 DE (Convertase PC5) (PC6) (PC6).
 GN PCKS5 OR PCS5 OR PCK5.
 OS Homo sapiens (Human).
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoide; Homo.
 OX NCBI_TaxID:9606; RN [1];
 RP SEQUENCE FROM N.A.
 RC TISSUE=T-cell;
 RX MEDLINE=6353880; PubMed=0755538;
 RA Miranda L., Wolf J., Pichantres S., Duke R., Franzusoff A.;
 RT "Isolation of the human PC6 gene encoding the putative host protease for HIV-1 gp160 processing in CD4+ T lymphocytes.,"
 RL Proc. Natl. Acad. Sci. U.S.A. 93:7695-7700(1996).
 RN [2];
 RP REVISIONS.
 RA Franzusoff A., Miranda L., Wolf J., Pichantres S., Lu Y., Duke R.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RN [3];
 SEQUENCE OF 15-913 FROM N.A.

RP Reudelhuber T.L.;
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: LIKELY TO REPRESENT A WIDESPREAD ENDOPROTEASE ACTIVITY WITHIN THE CONSTITUTIVE AND REGULATED SECRETORY PATHWAY. CAPABLE OF CLEAVAGE AT THE RXK(R) CONSENSUS MOTIF.
 CC -!- CATALYTIC ACTIVITY: RELEASE OF MATURE PROTEINS FROM THEIR PROTEINS BY CLEAVAGE OF ARG-XAA-YAA-ARG-|-ZAA BONDS, WHERE XAA CAN BE ANY AMINO ACID AND YAA IS ARG OR LYS.
 CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS SEEMS TO BE PRODUCED BY ALTERNATIVE SPLICING.
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN T-LYMPHOCYTES.
 CC -!- DOMAIN: THE PROPEPTIDE DOMAIN ACTS AS AN INTRAMOLECULAR CHAPERONE

CC ASSISTING THE FOLDING OF THE ZYMOGEN WITHIN THE ENDOPLASMIC RETICULUM; BELONGS TO PEPTIDASE FAMILY SB.
 CC -!- SIMILARITY: CONTAINS 1 HOMO B/P DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: U56387; AAC50643.2;
 DR EMBL; U9114; AAC91807.1; .;
 DR HSSP; Q99405; IMPR.
 DR MBROPS; S08.076; .;
 DR Genew; H9NC:8747; PCSKS.
 DR MM; 60088; .;
 DR InterPro; IPR02174; Furin-like.
 DR InterPro; IPR02884; P domain.
 DR InterPro; IPR000209; Peptidase_S8.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR Pfam; PF01483; P; PARTIAL.
 DR PRINTS; PR00723; SUBTILISIN.
 DR proDom; P000717; P-domain; 1.
 DR SMART; SM0261; FU; 5.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SBR; 1.
 KW Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal; Cleavage on pair of basic residues; Repeat; SIGNAL; 33 122 BY SIMILARITY.
 FT PROPEP 33 114 BY SIMILARITY.
 FT CHAIN 115 913 PROTEIN CONVERTASE SUBTILISIN/KEXIN
 FT DOMAIN 115 454 TYPE 5.
 FT DOMAIN 462 600 CYS-RICH MOTIF (CRM) REGION.
 FT DOMAIN 636 913 CLEAVAGE (AUTO-) (BY SIMILARITY).
 FT SITE 114 115 CELL ATTACHMENT SITE (POTENTIAL).
 FT SITE 519 521 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT SITE 171 171 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT SITE 212 212 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT SITE 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 225 225 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 381 381 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 665 665 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 752 752 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 802 802 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 852 852 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 118 118 S --> F (IN REF. 3).
 FT CONFLICT 121 121 V --> A (IN REF. 3).
 FT CONFLICT 511 511 R --> A (IN REF. 3).
 FT CONFLICT 601 601 R --> Q (IN REF. 3).
 FT CONFLICT 601 601 R --> Q (IN REF. 3).
 FT SEQUENCE 913 AA; 101775 MN; 21389264; CADP746C CRC64;

Query Match 10.7%; Score 162; DB 1; Length 913; Best Local Similarity 22.3%; Pred: 0.0002; Indels 78; Gaps 12; Matches 55; Conservative 33; Mismatches 81; InDel 78; Gaps 12;

OY 21 GSQNSRQRORRMPHNVNSQCGCATCSDYNG---CLSKPRLFALERIGHQIGVCL 78
 DB 677 GHYDAK-KRKRCAPN-----CSCFGSHGDDCGMSKYGYFL-----NEETNSCV 721

OY 79 SSQPSG-WYGRGPINKTKRQADQGCFNPKTCGSKFYIHLGKCLDNCPGLEAN 137
 DB 722 THEPDGSKQPTKK--NLRKCKSBNCKTCRTEGRLQGSGQCSVCEPGRVFN 778

OY 138 NHTME-----CVS-----VHCEVSEW-----NPWSCTKK 163
 DB 779 GQCPQPHRFATCAGAGADGCTCSCYFMDGRVQCSISVFDISSENYKSKKC 838

OY 164 GKTC-----GFKRGTE-----TRVREILOHPSAKGNLCPPTNETRKCTV 202

